



Vel Tech
Rangarajan Dr. Sagunthala
R&D Institute of Science and Technology
(Deemed to be University Estd. u/s 3 of UGC Act, 1956)

Avadi, Chennai, Tamil Nadu, India

**IEEE INTERNATIONAL
CONFERENCE**
ON
**ELECTRONIC SYSTEMS AND
INTELLIGENT COMPUTING**

April 22 – 23, 2022

**ICESIC
2022**

Sponsored by



New Delhi, India



IEEE
Madras Section



Table of Contents

<i>Message from Chief Patron Col. Prof. Dr. Vel. R. Rangarajan</i>	i
<i>Message from Chief Patron Dr. Mrs. Sagunthala Rangarajan</i>	ii
<i>Message from Chief Patron Mrs. Rangarajan Mahalakshmi Kishore</i>	iii
<i>Message from Patron Prof. Dr. S. Salivahanan</i>	iv
<i>Message from Dr V. Srinivasa Rao</i>	v
<i>Message General Co-Chair Dr M. Kavitha</i>	vi
<i>Message Conference Chair Dr V. Sivakumar</i>	vii
<i>Organizing Committee</i>	viii
<i>Technical Reviewer Board</i>	x
<i>International/National Advisory Committee</i>	xi
<i>About Vel Tech</i>	xii
<i>Keynote by Prof. Dr. Haklin Kimm</i>	xiv
<i>Keynote by Dr. Sarma Venkataraman</i>	xv
<i>Keynote by Dr.N. Kumarappan</i>	xv
<i>Keynote by Dr. Priyanka Kokil</i>	xvi
<i>Keynote by Dr Subhransu Sekhar Dash</i>	xvii
Towards Quantum Artificial Intelligence Electromagnetic Prediction Models for Ladder Logic Bombs and Faults in Programmable Logic Controllers <i>Varghese Mathew Vaidyan, Akhilesh Tyagi</i>	1
Deep Learning Approaches for Crack Detection in Bridge Concrete Structures <i>Daniel Einarson, Dawit Mengistu</i>	7
On-Demand Recharge Scheduling Algorithm in Wireless Sensor Networks <i>Bhargavi Dande, Chih-Yung Chang, Ming-Yang Su, Cheng-De Fan</i>	13
Mobile Charger Scheduling Algorithm for Energy Recharging in Wireless Sensor Networks <i>Bhargavi Dande, Chih-Yung Chang, Cheng-De Fan</i>	18
Utilization of Fog Computing in Task Scheduling and Offloading: Modern Growth and Future Challenges <i>Archana R.</i>	23
Design of Deep Learning Based Covid 19 Diagnosis Framework Using Lung Ultrasound Images <i>Rajesh Doss</i>	35
Artificial Intelligence Model for Classification of Sensitivity Data Utilized At Industrial Applications <i>Ashraf Mohammad, Mupparaju Kavya Sree, Saam Prasanth Dheeraj Pedapalli, Muddineni Raveen, Yaswanth Pavan Kumar. T</i>	41
Smart V2G/G2V Charging for Grid Connected Electric Vehicles in Indian Scenario <i>Setlem Bhargavi, Uppuluri Baby Jyothsna, K. Deepa, V. Sailaja, J. Ramprabhakar</i>	46
Categorization of Fruit Images Using Artificial Bee Colony Algorithm Based On Glcm Features <i>Akshay S, Deepika Shetty M A</i>	52
Detection of Falsified Selfish Node with Optimized Trust Computation Model in Chimp–AODV Based Wsn <i>P. Manjula, Sbaghavathi Priya</i>	58

An Improved Linguistic Haar Fuzzy Decision Maps <i>Karthik S, A Felix, Selvaraj A, Gunasekar T</i>	64
Classification of Pulmonary Emphysema Using Deep Learning <i>R. Ramadoss, C. Vimala</i>	69
Lung Cancer Detection Using Vgg-16 Deep Learning Model <i>K. Ramanjaneyulu, K. Hemanth Kumar, K. Snehith, G. Jyothirmai, K. Venkata Krishna</i>	75
Credit Card Fraud Detection Using Machine Learning <i>Janet B</i>	79
An Anomaly Based Network Intrusion Detection System Using Lstm and Gru <i>Rachana Koniki, Mounika Durga Ampapurapu, Praveen Kumar Kollu</i>	85
Diagnosing Musculoskeletal Disorders From Shoulder Radiographs Using Deep Learning Models <i>A. Kavitha, T. Kujani, S. Gokul Kannan, V. Akila</i>	91
Piezoelectric Transducer Power Generation for Low Power Applications <i>B Rubini, S Manoj</i>	97
Coronary Artery Disease Prediction Using Machine Learning Algorithm <i>B Baranidharan, Kaushik Kumaran, Adhith Sankar, Keshav Rao</i>	101
Estimation of Soil Salination for Crop -Yielding and How to Prevent <i>Reshma Neppalli, Ch. Gowthami Supriya, A. Hema, K. Kiran Kumar, N. V. K Ramesh, B. Naresh Kumar Reddy</i>	107
Transmission of Biometric Feature Using Facial Features Securely for Long Distance Biometric Recognition System <i>Suganthy M, S Manjula, M Kavitha, P Anandan</i>	111
Real Time Continuous Behavior Class Based Secure Routing for Improved Qos in Industrial Networks <i>Babu. N, Tamilarasi Suresh</i>	117
Opinion Mining Models for Learner Feedback On Massive Open Online Courses <i>Christy Jackson J, Naren G O, Suriyakrishnan Sathish, Maheysh V, Aravinda Boovaraghavan, Makesh Srinivasan, Hashwanth S</i>	123
Automatic Stress Recognition System with Deep Learning Using Multimodal Psychological Data <i>Praveenkumar S, Karthick T</i>	128
Performance of Indian Cricket Team in Test Cricket: A Comprehensive Data Science Analysis <i>Vishnusai Viswajith Tharoor</i>	134
Deep Learning Based Classification of Cervical Cancer Using Transfer Learning <i>Hemalatha K, V Vetriselvi</i>	140
Performance Model Evaluation of Seizure Classification Using Statistical Features and Random Forest <i>S. Ramya, M. Uma</i>	146
An Efficient Key Agreement and Anonymous Mutual Authentication Protocols for Secure Communication in Vanets <i>Swathi Kona, Sai Venkata Krishna Reddy Morthala, Rakesh Konathala, Pawan Kumar Pinninti, Hemanth Kumar Mavuru, Azees Maria</i>	152
A Robust Approach Based On Cnn-Lstm Network for the Identification of Diabetic Retinopathy From Fundus Images <i>Sajeev Ram Arumugam, E. Anna Devi, V. Rajeshram, Balakrishna R, Sankar Ganesh Karuppasamy, S. Vinod Kumar</i>	158
Telefit- an Iot Enabled Health Care Device <i>Soma Prathibha, Shyamkumar M, Kirthiga M, Srinidhi S, Yashwanth Krishnan B, Pravin Kumar P, Shabika Fathima M</i>	163
Iot Based Collaborative Industrial Gas Prediction and Warning Using Machine Machine Learning <i>B R Sathishkumar : S Allirani : M Sree Dhviya : V Madhumithaa</i>	168

A Comparative Approach for Investigating an Efficient Method for Reduction of Search Space in Lung Image Analysis <i>Sheetal Umakant Pawar</i>	173
Private Messaging Service Using AES Encryption and Toxicity Detection <i>Shanmuga Priya R, Tharunraj M, Rozen Berg D, Sanjay Kannan M</i>	179
Application of Ann, Svm and Knn in the Prediction of Diabetes Mellitus <i>M Revathi, A Beena Godbin, S Nikkath Bushra, S Anslam Sibi</i>	185
Classification of Ecg Signal Using Machine Learning Algorithm <i>Md Tarique Anis, Veena Sharma</i>	191
Path Planning of Intelligent Uavs Using Computational Geometric Techniques <i>Akshya J, P. L. K. Priyadarsini, Prudhvi Gadupudi, Karthik Paladugula</i>	196
A Noval Multi Carrier Pwm Based Harmonic Mitigation of Multilevel Pv Inverter with Grid Integration <i>Satyanarayana. Vanam, Manne Navya Sree</i>	200
Power Quantity Improvement in Grid System with Pv Based Svpwm-Dvr for Sag & Swell Mitigation with A Novel Ann Controller <i>Satyanarayan. Vanam, Banovath Shiva Prasad Bharath</i>	206
Hybrid Energy Management System & Power Compensation of PV & Wind Integrated Grid System with Fuzzy Based D-Statcom <i>Satyanarayana. Vanam, Gurijalasreedhar</i>	212
Densenet201 for Animal Detection and Repellent System <i>Naveenkuamr M, Manoj R, Nandhakumar B, Rahul R</i>	219
An Optimized Gaussian Extreme Learning Machine (Gelm) for Predicting the Crop Yield Using Soil Factors <i>R. Prabavathi, Balika J Chelliah</i>	225
Maneuvering Control of Autonomous Underwater Vehicle (Auv) By Simulated Annealing and Moth Flame Optimization Tuned Controllers <i>Sudarshan K. Valluru, Hitesh Thareja</i>	229
Design and Realization of Wideband 4- Element Printed Mimo Antenna for X -Ku- Band Satellite and Limited K- Band Short Range Applications <i>Ch Muralikrishna, Er Jeetamitra Satapathy, N Suguna, Sunita Singh, Ganta Sandeep V Padmakar, Rayudu Bharath Kumar</i>	235
Artificial Intelligence Based Real Time Lie Detector Using Eye Gaze Pattern <i>Krishnakishore K, Durga Kameswari S, Jagannaveen V</i>	241
Eb Algorithm for Effective Privacy and Security of Data Processing in Mcc <i>V. S. Prakash, N. Bharathiraja, R. Deiva Nayagam, R. Thiagarajan, R. Krishnamoorthy, J. Omana</i>	247
Hybrid Filtee and Wrapper Method Based Feature Selection for Crop Recommendation <i>Madhuri J, Indiramma M</i>	253
Portable Time Governing Switchbox Using Internet of Things <i>Chansekhar S, Manikandan A, Dinesh P, Logesh S, Hari Prasath B</i>	259
Toxic Speech Classification Using Machine Learning Algorithms <i>Pabba Sumanth, Syed Samiuddin, K. Jamal, Srikanth Domakon, Pathi Shivani</i>	263
Implementation of Ann for Predication of Performance and Emission Characteristics in Ci Engine <i>Narennathan. S. K, M. Thiyyagu, Mr. Shankar. A, Mr. T. N. Sureshkumar</i>	270

Design of an All-Digital Phase-Locked Loop in a 130Nm Cmos Process Using Open-Source Tools <i>Charaan S, Nalinkumar S, Elavarasan P, Prakash P, Kasthuri P</i>	276
A Term Weight Measure Based Approach for Author Profiling <i>Karunakar Kavuri, M Kavitha</i>	281
Joint Spectral-Spatial Feature Using Deep 3-D Cnn for Hyperspectral Images <i>M Kavitha, R Gayathri</i>	287
Traffic Violation Invigilation Using Transfer Learning <i>N R Rajalakshmi, K Saravanan</i>	292
Performance Evaluation of Convolution Neural Network for Lung Cancer Detection <i>Dhasny Lydia M, Prakash M</i>	299
Decision Trees to Detect Malware in A Cloud Computing Environment <i>N Vijayaraj, M Sumathy, Rajkamal M ,Uganya,T Kamaleshwar,D Rajalakshmi</i>	305
Evaluating Performance Metrics in Classifying Bitcoin Mixing Services Using Decision Tree Algorithm <i>Ezhilmathi S, Selvakumara Samy S</i>	310
Memristor Based Cam Cell Designs and Analysis of Their Performance <i>Jugi Rithanya Rp, Pooja K, Sakthivel R</i>	317
Locational Marginal Pricing Based Management of Congestion with Optimum Sizing of Distributed Generator Using Modified Ilshade Algorithm <i>Prashant, Anwar Shhzad Siddiqui, Md. Sarwar</i>	323
Blur Classification and Estimation of Motion Blur Parameters Using Olr <i>Sakthivel S, Akash Ram R. K, Sidarth Sai B, Ganesh Kumar C</i>	329
Simulation of Machine Learning Techniques to Predict Academic Performance <i>Praveena Chakrapani, Chitra Devi D</i>	335
Academic Performance Prediction Using Machine Learning: A Comprehensive & Systematic Review <i>Praveena Chakrapani, Chitra Devi D</i>	341
Rhythm Monitor - A Wearable for Circadian Health Monitoring <i>K. Sornalakshmi, Revathi Venkataraman, N. Shalin, M. Jerome Samraj, M. Viveka</i>	347
Identification of Voice Pathology From Temporal and Cepstral Features for Vowel 'a'Low Intonation <i>Gayathri S, Priya E</i>	351
Smart Video Surveillance Based Weapon Identification Using Yolo V5 <i>S. Nikkath Bushra, G. Shobana, K. Uma Maheshwari, Nalini Subramaniam</i>	357
Parameter Tuned Unsupervised Fuzzy Deep Learning for Clinical Data Classification <i>S. S. Saranya, N. Sabiyath Fatima</i>	364
Competition Prediction and Analysis Using Machine Learning <i>Vijay K, K. Jayashree, Vijayakumar R,Babu Rajenn</i>	370
Integrated Communal Attentive & Warning System Via Cellular Systems <i>Vijayakumar R, Ravikumar S, Vijay K, Sivarajanji P</i>	376
Real-Time Smart Attendance Monitoring System with Thermal Scanning <i>Gorla Sujan Souris, Gummala Sainath Reddy, Panditi Premsagar, V. Hima Deepthi</i>	382
Renewable Energy Sources and Electric Vehicle for Optimal Energy Management of Micro Grids with the Aim of Co2 Emission Reduction <i>S. Naveen Prakash, N. Kumarappan</i>	388
<i>Author Index</i>	393

Preface

We feel greatly honoured to have been assigned the job of organising the IEEE International Conference on Electronic Systems and Intelligent Computing (ICESIC 2022) on April 22nd and 23rd, 2022, at Veltech Rangarajan Dr Sagunthala R & D Institute of Science and Technology.

It is our pleasure and privilege to present to you the proceedings of the International Conference in a bound volume, for the benefit of the participants and others. The main objective of the conference is to bring together researchers, scientists, engineers, and research scholars in one place to discuss the challenges encountered and the solutions adopted in the fields of Artificial Intelligence and Electronic expert systems in engineering. It is about 68 papers that have been accepted out of 572 papers received from various countries and have been extremely encouraging. The conference has been divided into major sessions like AI and Machine Learning, Computer Vision and Deep Learning, IoT and Wireless Communication, and Electronics, Devices, Circuits and Systems.

We take this opportunity to thank our honourable Founder Chancellor & President, Col.Prof.Dr. Vel. R. Rangarajan, and Founderess, Dr. Sagunthala Rangarajan, wholeheartedly for the confidence they entrusted in us to organise this International conference, ICESIC-2022. We sincerely thank our beloved Chairperson & Managing Trustee, Mrs. Rangarajan Mahalakshmi Kishore for her continuous support and guidance in organising the conference. We would like to express our heartfelt gratitude to the Vice Chancellor Prof. Dr. S. Salivahanan and Dean SoC Prof. Dr V. Srinivasa Rao of Veltech Rangarajan Dr Sagunthala R & D Institute of Science and Technology for their invaluable assistance in making this programme a success.

We thank all the authors of the papers and participants in the conference for making this event a grand success. We thank our sponsors, IEEE Madras Section and AICTE, for their technical and financial support of our conference.

All the accepted and registered papers will be submitted for inclusion into IEEE Xplore (Scopus Indexed).

Welcome Message from Chief Patron



Col.Prof. Dr. Vel. R. Rangarajan
Founder Chancellor & President
Chief Patron

Dear Delegates,

I have a great pleasure in inviting you all to the 2022-IEEE International Conference on Electronic Systems and Intelligent Computing (ICESIC) being held at Veltech Rangarajan Dr Sagunthala R&D Institute of Science and Technology.

I congratulate the Department of Computer Science and Engineering for organizing this International Conference. The theme of the conference are appropriate to the present day needs of the technology. I hope the conference will also deliver the pathways and to explore the scientific and technological know-how. I wish to place on record my sincere appreciation for the excellent co-operation extended by the IEEE Madras Section and AICTE for Co-sponsoring this conference.

I am confident that every participant will carry pleasant memories of this event. I wish the conference all success.

A handwritten signature in blue ink, appearing to read "B.R.R." followed by a surname.

(Col.Prof. Dr. Vel. R. Rangarajan)

Welcome Message from Chief Patron(s)



Dr. Mrs. Sagunthala Rangarajan
*Founderess President
Chief Patron*

Dear Delegates,

I feel ecstatic to know that the Department of Computer Science and Engineering is organizing IEEE International Conference on Electronic Systems and Intelligent Computing (ICESIC) on April 22 & 23, 2022. I am sure that this International conference will provide an opportunity to the Scientist, Researchers, Engineers, Expert from Industry and students to share their innovative ideas in their recent findings, problems and solutions. I also expect that this conference will be a platform for the benefit of present and future technology.

I wish all the delegates of the conference very best, and whole heartedly congratulate all the faculty members for making this conference a great success.

A handwritten signature in blue ink, appearing to read "Dr. Mrs. Sagunthala Rangarajan".

(Dr. Mrs. Sagunthala Rangarajan)



**Mrs. Rangarajan Mahalakshmi
Kishore**
*Chairperson & Managing Trustee
Chief Patron*

Dear Delegates,

I am glad to know the commendable initiative take by the Department of Computer Science & Engineering, to conduct the IEEE International Conference on Electronic Systems and Intelligent Computing (ICESIC) on 22nd April and 23rd April 2022. ICESIC-22 provides a forum for the dissemination of new ideas, research and development, practical experiments with its results and also concentrating on both theory and practices of intelligent systems.

I wish the conference to become an inspiration for the development and successful adaption of smart intelligent technologies in India and abroad.



(Mrs. Rangarajan Mahalakshmi Kishore)

Welcome Message from Patron



Prof. Dr. S. Salivahanan
Vice Chancellor
Patron

Dear Delegates,

It gives a sense of immense pleasure to know that the Department of Computer Science and Engineering is organizing the "IEEE International conference on Electronic Systems and Intelligent Computing" (ICESIC-22) on April 22 & 23, 2022.

Computers are becoming smarter, as artificial intelligence and machine learning are making tremendous strides in simulating human thinking. Creating computer systems that automatically improve with experience has many applications including robotic control, data mining, autonomous navigation and bioinformatics. Most AI projects contribute in reducing costs, increasing revenue, launching new line of business.

I believe that this conference will provide a platform to share a wide knowledge, explore more of application oriented concepts and provide a rich experience to all its participants. I wish the conference a grand success

A handwritten signature in blue ink that reads "S. Salivahanan".

(Prof. Dr. S. Salivahanan)

Welcome Message from General Chair



Dr V. Srinivasa Rao
*Dean School of Computing
General Chair*

Dear Delegates,

I have great pleasure to welcome you all to the IEEE International conference on Electronic Systems and Intelligent Computing (ICESIC-22). I am very much delighted to note that the overwhelming response for our invitation from authors, research scholars and Industry people for this conference. The technical program committee and reviewers worked with excellence in selecting high quality papers for oral presentation in the conference and inclusion in the IEEE Explore.

Intelligent systems is poised to transform every industry just as electricity did 100 years ago. It will create \$13 trillion of GDP growth by 2030 according to MC Kinsey, most of which are not-internet sectors including manufacturing, agriculture, energy, logistic and education. This conference encourages the youthful potential Engineering brains from all over the country, to gather and emerge with constructive ideas towards technical excellence.

I believe that this conference will provide a platform to share a wide knowledge, explore more of application oriented concepts and provide a rich experience to all its participants.

I wish the conference a grand success.

A handwritten signature in black ink, appearing to read "L - 2".

(Dr. V. Srinivasa Rao)

Welcome Message from General Co-Chair



Dr M. Kavitha
Professor
General Co-Chair

Dear Delegates,

I am indeed most delighted to be General Co-Chair of the AICTE and IEEE Sponsored International Conference on "Electronic Systems and Intelligent Computing (ICESIC-2022)" being organized on April 22nd and 23rd, 2022 by the Department of CSE, School of Computing, Vel Tech Rangarajan Dr Sagunthala R & D Institute of Science and Technology .

The topics of the conference are found in conformity with Gartner's top strategic predictions for 2022 and beyond. The computer was born out of a need to solve a serious number-crunch crisis. Nowadays, number crunching has become trivial as we carry more computing power on our smartphones than was available in early models. A human being can perceive, learn, think, decide, and act, which are essential traits for enabling adaptive behaviour in a dynamic environment. The ultimate objective is to enable the machine to mimic human beings in decision making and acting, adapting to the traits of perception, learning, and thinking.

Essentially, the topics of the conference reflect emerging thinking on how to make a machine learn so that it can mimic human beings. The Department of CSE deserves appreciation for organizing this conference at the right juncture.

I hope that the authors of papers and participants will return to their places with valuable take-home messages that act as motivators for pursuing research in emerging areas.

I wish the conference a grand success.

A handwritten signature in blue ink that appears to read "Kavitha".

(Dr. M. Kavitha)

Welcome Message from Conference Chair



Dr V. Sivakumar
Associate Professor
Conference Chair

Dear Delegates,

It is my great honour to welcome you all the delegates for the IEEE International conference on Electronic Systems and Intelligent Computing (ICESIC-22). The researchers shall share their research outcomes with peer groups by presenting the same in conferences and / or publishing in reputed journals. This is to provide scope for further advancements in knowledge and innovations thereof. The ICESIC would facilitate its participants for sharing their research outcomes with peer groups.

It is sure that the conference proceedings would be significant for furthering research in the interdisciplinary areas such as intelligent things and intelligent electronics applications.

Eventually, I would like to thank our Dean-SOC and Management for their guidance, support and motivation for the success of this International Conference.

A handwritten signature in blue ink, appearing to read "V. Sivakumar".

(Dr. V. Sivakumar)

Organizing Committee

Chief Patrons

Col. Prof. Dr. Vel. R. Rangarajan, *Chancellor & Founder President.*
Dr. Sagunthala Rangarajan, *Foundress President.*
Mrs. Rangarajan Mahalakshmi Kishore, *Chairperson & Managing Trustee.*

Patrons

Prof. Dr. S. Salivahanan, Vice Chancellor.
Prof. Dr. E. Kannan, Registrar.

General Chairs

Dr V. Srinivasa Rao, Vel Tech, Chennai, India.
Dr N. Kumarappan, Chairman IEEE Madras Section, India.

General Co-Chair

Dr. M.Kavitha, Vel Tech, Chennai, India.
Dr. Subhransu Sekhar Dash,GCE,Keonjhar, India.
Dr. Ramazan Bayindir, Gazi University, Turkey.

Conference Chair

Dr. V. Sivakumar, Vel Tech, Chennai.

Organizing Chair

Dr V. Srinivasa Rao, Vel Tech, Chennai.

Program Chairs

Dr.M.Kavitha, Vel Tech, Chennai.
Dr.N.R.Rajalakshmi, Vel Tech, Chennai.

Program Co-Chairs

Dr.N.Gomathi, Vel Tech, Chennai.
Dr. Preetisudha Meher, NIT Arunachal Pradesh.
Dr. R. Ganesan, VIT, Chennai.

Keynote Speakers

Dr. Haklin Kimm, East Stroudsburg University, USA.
Dr. Sarma Venkataraman, Director, CR Rao AIMSCS, India.
Dr. N. Kumarappan, Annamalai University, Tamil Nadu.
Dr. Priyanka Kokil, Indian Institute of Information Technology, Design & Manufacturing, India.
Dr. Subhransu Sekhar Dash, Government College of Engineering, India.

Special Session Chair

Dr.P.N. Suganthan, NTU Singapore.
Dr.R.Sakkavarthi, Vanier College, Quebec, Canada.

Publicity Chair

Dr.V.Dhilip Kumar, Vel Tech, Chennai.
Dr.Padmanaban, Vel Tech, Chennai.

Publication Chair

Dr.R.Kavitha, Vel Tech, Chennai.
Dr.M.Shyamala Devi, Vel Tech, Chennai.
Dr.L.Thanga Marriappan, Vel Tech, Chennai.
Dr.J.Arunpandian, Vel Tech, Chennai.

Finance Chair

Dr.Carmel Mary Belinda, Vel Tech, Chennai.
Mr.S.Ravikumar, Vel Tech, Chennai.

Organizing Committee Members

Dr.N.Malarvizhi, Vel Tech, Chennai.
Dr.R.Srinivasan, Vel Tech, Chennai.
Dr.J.Visumathi, Vel Tech, Chennai.
Dr.G.R. Kanagachidambaresan, Vel Tech, Chennai.
Dr.D.Umanandhini, Vel Tech, Chennai.
Dr.S.Sankara Narayanan, Vel Tech, Chennai.
Dr.N.Vijayaraj, Vel Tech, Chennai, India.
Dr.R.Aruna , Vel Tech, Chennai.
Dr.N.Karthik, Vel Tech, Chennai.
Dr.G.Tamilmani, Vel Tech, Chennai.
Mr.R.Ganesan, Vel Tech, Chennai.
Mr.M.Thanjai Vadivel, Vel Tech, Chennai, India.
Mr.D.Manivannan, Vel Tech, Chennai, India.
Mr.N.K. Manikandan, Vel Tech, Chennai, India.
Mr.K.Kishore Kumar, Vel Tech, Chennai.
Mr.K.Antony kumar, Vel Tech, Chennai.
Mr.S.Durai, Vel Tech, Chennai.
Mrs.K.Prema, Vel Tech, Chennai.
Mrs.T.Kujani, Vel Tech, Chennai.
Mr.Ignatious K Pious, Vel Tech, Chennai.
Mr.J.Vivek, Vel Tech, Chennai.
Mrs.P.Arivubrakan, Vel Tech, Chennai, India.

Technical Reviewer Board

Dr. Selwyn Piramuthu, University of Florida, USA.
Dr. Yvan Labiche, Carleton University, Ottawa, Canada.
Dr. Amina Omrane, University of Sfax in Tunisia.
Dr. Rajasekar Ramalingam, MHD, Oman.
Dr. Yuvaraja Teekaraman, Vrije Universiteit Brussel.
Dr. Subramaniam Ganesan, Oakland University, USA.
Dr. Meenalosini Vimal Cruz , GSU, Georgia.
Dr. Po-Ming Lee, STUST, Tiwan.
Dr. Basim Alhadidi, AlBalqa' Applied University, Jordan.
Dr. Kasi Periyasamy, University of Wisconsin-La Crosse, USA.
Dr. Daniel Einarson, Kristianstad University, Sweden.
Dr. Antonis Sidiropoulos, ATET, Greece.
Dr. Geetha M, University of West London, UAE.
Dr. Naveen Chilamkurti, La Trobe University, Australia.
Dr. Jeyaprakash Chelladurai, ESU, Pennsylvania.
Dr. Suzanne Kieffer, Université Catholique de Louvain,Belgium.
Dr. Goran Dambic, Algebra University College,Croatia.
Dr. Keshav Dahal, University of West of Scotland,UK.
Dr Korhan Cengiz,Trakya University,Turkey.
Dr. Kemal Polat, Bolu Abant Izzet Baysal University,Turkey.
Dr. Amina Omrane, University of Sfax, Tunisia.
Dr. Nadia Mansour, University of sousse Tunisia, Spain.
Dr. Mohamed Lahby, UHC, Morocco.
Dr. Jagabar Sathik M, Prince Sultan University, Saudi Arabia.
Dr. Asutosh Kar, IIITDM Kancheepuram.
Dr. Nedumaran Damodaran, University of Madras.
Dr. R. Ramaprabha, SSN College of Engineering.
Dr. S Radha, SSN College of Engineering.
Dr. N. Venkateswaran, SSN College of Engineering.
Dr. Sudha, SASTRA deemed university.
Dr. Kaliraj, Manipal Academy of Higher Education, Manipal.
Dr. M.Senthamil Selvi, Sri Ramakrishna Engineering College.
Dr. Muthukumaran, Hindustan IST, Chennai.
Dr. Sridhar ,SRM IST, Chennai.
Dr. Rajkumar, VIT Chennai.
Dr. Magesh Kasthuri, Wipro, Bengaluru.
Dr. Surekha ,Amrita School of Engineering, Bengaluru.
Dr. S.Albert Alexander, Kongu Engineering College.
Dr. P. Sivakumar, Dr NGP Institute of Technology, Coimbatore.
Dr.K.Rajalakshmi, PSG College of Technology, Coimbatore.

International/National Advisory Committee

Dr. M.Sridhar, University of Michigan, Dearborn, USA.
Dr. Vijayan Sugumaran, Oakland University, USA.
Dr. Mario Molinara, University of Cassino, Italy.
Dr. Mohamed Elhoseny, AUE, Dubai, UAE.
Dr. Xiao-Zhi Gao, University of Eastern Finland, Finland.
Dr. Thinagaran Perumal, University Putra Malaysia, Malaysia.
Dr. Haklin Kimm, East Stroudsburg University, USA.
Dr. Fazirulhisyam Hashim, Universiti Putra Malaysia, KL.
Dr. Ahmed Elnagar, Beni-Suef University, Egypt.
Dr. R. Logeswaran, Asia Pacific University, Malaysia.
Dr. L.J.Yang, Tamkang University, Taiwan.
Dr. Muhammad Shafiq, Guangzhou University, China.
Dr. Vijayakumar Varadarajan, UNSW, Sydney, Australia.
Dr. Anand Paul, Kyungpook National University, Korea.
Dr. Oguz Bayat, Altınlba_University, Turkey.
Dr. Oana Geman, University of Suceava, Romania
Dr. D. Devaraj Secretary, IEEE Madras Section
Dr. S. Joseph Gladwin, Treasurer, IEEE Madras Section
Dr. E. Kannan, Vel Tech, Chennai.
Dr. S. Koteswaran, Vel Tech, Chennai, India.
Dr. P. Chandrakumar, Vel Tech, Chennai.
Dr. A.T. Ravichandran, Vel Tech, Chennai.
Dr. R. Suguna, Vel Tech, Chennai.
Dr. Rajendra Prasath, IIIT, Sri City, Chittoor.
Dr. Victer Paul, IIIT Kottayam, Kerala.
Dr. Masilamani, IIIT, Kanchipuram.
Dr. Srinivasa Krishna Srivatsa, MIT, Chennai.
Dr. A John Prakash, Anna University, Chennai.
Dr. P. Yogesh, Anna University, Chennai.
Dr. D. Ezhilarasi, NIT Trichy.
Dr. Bhavna Arora, Central University of Jammu.
Dr. A. Martin, Central University of Tamil Nadu.
Dr. T S Pradeep Kumar, VIT Chennai.
Dr. V S Kanchana Bhaaskaran VIT Chennai.
Dr. Debdatta Kandar, North-Eastern Hill University, Shillong
Dr. Revathi Venkataraman, SRMIST, Chennai.
Dr. C. Lakshmi, SRMIST, Chennai.
Dr. B. Amutha, SRMIST, Chennai.
Dr. S. Karthik, SRMIST, Chennai.
Dr. P. Ambika, Expert Data Scientist, Bangalore.
Dr. Utpal Roy, Visva-Bharati University, West Bengal.
Dr. Christy Jackson, VIT Chennai.
Dr. Muhammad Abulaish, South Asian University, Delhi.

About Vel Tech

About Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology

Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology (Vel Tech) is one of the pioneering Deemed Universities of INDIA. It was established in 1997. It offers a wide range of programs, including undergraduate, post-graduate, and doctoral programs. Vel Tech is the first member from India to adopt Conceive-Design-Implement-Operate (CDIO). CDIO is a worldwide framework which aims to produce the next generation of engineers. The CDIO based academic process undoubtedly generates knowledgeable, up-skilled graduates with a professional attitude who are industry ready.

Global Alliances of Vel Tech has working International relations with more than 390 Institutions across 30 countries and has entered into a Memorandum of Understanding (MoU) with 120 institutions for various academic and research exchange activities. Students, faculty members, and researchers get a collaborative platform for academic projects, explore leading best practices, and research work with International Universities.



Vel Tech believes that there are no boundaries in the pursuit of knowledge. Through these strategic alliances with International Universities, the standard of education at Vel Tech will be on par with the best in the world. Vel Tech holds a National Institute Ranking Framework (NIRF) ranking, Atal Ranking of Institutions on Innovation Achievements (ARIIA) by following the parameters that broadly cover "Teaching, Learning and Resources," "Research and Professional Practices," "Graduation Outcomes," "Outreach and Inclusivity," and "Perception".

Vel Tech has created an ecosystem for innovations, including an Incubation Centre and other initiatives for the creation and transfer of knowledge. The whole initiative supports early stage startups by providing incubation space, mentoring, networking, & financial loans (Convertible debt) up to INR 50.00 Lakhs. Vel Tech Technology Business Incubator (TBI) has created state of the art facilities like NVIDIA supported AI, ML, DL & digital services labs to support IT innovations. Vel Tech TBI is recognized as a Centre of Excellence under the NIDHI Scheme to expand the facility to a global level.

About Department of Computer Science & Engineering



Computer Science & Engineering (CSE) is one of the significant programs in the School of Computing. CSE enhances students' ability to develop & sustain the transformative software systems of tomorrow. The department activities are determined to create future leaders & trend setters for the next generation.

This department offers a comprehensive Undergraduate & Post Graduate programs curriculum that prepares students to be leaders in emerging technologies. Well qualified and experienced faculty form the backbone of the department. Experts from various Industries are invited to deliver class room sessions, Industry based hands-on activities, guest lectures and seminars related to the curriculum and software industry trending topics. Currently, B.Tech. Computer Science & Engineering program is accredited by NBA under the Tier-I (Washington Accord) category.

Keynote Speakers



Prof. Dr. Haklin Kimm
Professor of Computer Science

B.S., 1979, Korea University at Seoul
M.S., 1984, University of Oklahoma
Ph.D., 1988, University of Oklahoma

Email : hkimm@esu.edu; Phone: (570) 422-3523

Haklin Kimm is a Professor of Computer Science at East Stroudsburg University where he leads the High Performance Computing Lab. Kimm's research interests are in Distributed and Parallel Systems, Ubiquitous Computing and Communications, Automotive Software Engineering, Multilevel Security with Security enhanced Linux.

Teaching Interest

Mobile Computing (Grad) Embedded Systems (Grad)
Computer Networks (Grad/Under) Computer Architecture (Grad/Under)
Distributed Object Programming (Grad /Under) Advanced Operating Systems (Grad)
Operating Systems using Linux (Under) Probability for Computer Science (Under)
Object-Oriented Programming Parallel Computing using PVM/MPI
Assembly Language using IBM / PC Database Systems
Artificial Intelligence using Prolog/Neural nets Data Structure and Algorithms
Programming Language and Theory Network and Operating System Security
Cryptographic System Development Introduction to Computer Security

Research Interest

GPS Tracking Algorithms using GPS and Accelerometer, Application of OpenMP/MPI and Google TensorFlow programming, Distributed Image forgery detection algorithm, Starvation-free Controller Area Network Protocol

Service

ABET Computing Accreditation Commission Program Evaluator;
Invited Reviewer for Natural Sciences and Engineering Research Council of CANADA;
External PhD Thesis Reviewer for Brunei University;
Editorial Board Member for Journal of Advances in Computer Networks and Journal of Computer Science Applications and Information Technology.



SARMA VENKATARAMAN

Currently Director, CR Rao AIMSCS Institute in University of

Hyderabad, Hyderabad.

Retired from ISRO as Scientist H, Group Director

S. Venkataraman is working at ADRIN, Dept. of Space as a Group Director in the area of Systems Engineering Group. He has worked at Defence Electronics Research Laboratory for 5 years.

He has worked at George Tech, Atlanta, USA in DOD project in Artificial Intelligence techniques for decision making.

At ADRIN (Department of Space), he has been responsible in development of Artificial Intelligence techniques in Satellite Image processing and Analysis. He was heading a group of scientists in the area of cyber security solutions for data and network security. Software and Embedded based security solutions have been developed for various projects. He has 34 years of work experience and more than 65 publications in International Journals.

Currently he is working on development of a Block chain platform and Deep Learning methods for Real time Satellite Imagery classifications.



Dr. N. Kumarappan
Chairman, IEEE Madras Section,
Professor
Department of Electrical Engineering
Annamalai University
Tamil Nadu

Dr. N . Kumarappan received the Graduate degree from Madurai Kamaraj University, Tamil Nadu, India in 1982, the Post-Graduate degree from Annamalai University, Annamalai Nagar, India in 1989 and the Ph.D. degree from CEG Anna University, Tamil Nadu, India in 2004 under QIP fellowship AICTE, India. He is the former head of the department and currently a Professor with the Department of Electrical Engineering his citation is 1185, H index 18. He was the outstanding reviewer for the Elsevier, Faculty of Engineering and Technology, Annamalai University. He is having 34 years of experience to his credit as an educator and researcher. He has published more than 132 international journal and conference papers. According to Google scholar international journal of electric power and energy system 2015. He is an IEEE Madras Section Chairman, IEEE CIS Madras Chapter Chair and a Coordinator for more than 200 IEEE Madras Section organized FDP, Workshop, tech meet and TISP programs ect.. He was the recipient of the IEEE-NNS Outstanding Paper Travel Grant Award, Australia 2002, IEEE-WCCI Outstanding Paper Travel Grant Award, Canada 2016, the IEEE PES Student Program Award, USA 2003 and IEEE CIS Madras Chapter Best Chapter Award 2017. He was also a recipient of best researcher award, Annamalai University, 2018. His interview appeared in Electrical India magazine in February 2021 issue. He has acted as a guest editor in the special issue of Complexity journal, published by Wiley - Hindawi. He is the General Chair for the international conference MASCON 2021 organised by IEEE Madras Section. His current research interests include power system operation and control, electricity price forecasting, EHV transmission fault diagnosis, FACTS devices, power system reliability, artificial intelligence techniques, micro grid, distributed generation and smart grid. Dr.Kumarappan is a Life Fellow of the Institution of Engineer's (India) and a Life Member of the Indian Society of Technical Education. He was the recipient of the IEEE Madras Section Motivation Award in 2010 and the Certificate of Appreciation in 2012, third rank in the chapter activities 2015, CIS Madras Chapter Best Chapter Award 2021. He was the organizer of a special session in IEEE WCCI 2010 Spain, IEEE WCCI 2016 Canada and IEEE WCCI 2018 Brazil. He is a Speaker, Reviewer and Session Chair for IEEE/IET-U.K. international conferences and other international journals.



Dr. Priyanka Kokil
Assistant Professor, Head of the Department (ECE)
E-mail: priyanka@iiitdm.ac.in | Ph: +91-44-27476360

Personal: <https://bit.ly/3s5BOxa>
LinkedIn: <https://www.linkedin.com/feed/>
ResearchGate : <https://www.researchgate.net/profile/Priyanka-Kokil>

Education

Motilal Nehru National Institute of Technology, Allahabad, Allahabad

Ph. D. in Electronics Engineering, 2010 - 2013

IIT Delhi, Delhi

M. Tech. in Telecommunication Technology and Management, 2005 - 2007

Motilal Nehru National Institute of Technology, Allahabad

B. Tech. in Electronics Engineering, 1999 - 2003

Specialization

Digital Signal Processing

Biomedical Image Processing

Research Interests

System theory, Nonlinear systems, Biomedical Signal Processing, Machine Learning, Deep Learning, Multi-dimensional Signal Processing, Time delayed systems, Robust stability, System theory



Dr Subhransu Sekhar Dash
Professor and HoD
Department Electrical Engineering
Email_ID Subhransudash_fee@gcekjr.ac.in: Phone No: 9884356933

Dr. Subhransu Sekhar Dash is presently a Professor and HOD in the Department of Electrical Engineering, Government College of Engineering, Keonjhar, Odisha, India. He received his Ph.D. degree from College of Engineering, Guindy, Anna University, Chennai, India. He has more than 22 years of research and teaching experience. His research areas are AI techniques application to power system, modeling of FACTS controller, power quality and smart grid. He is a Visiting Professor at Francois Rabelais University, POLYTECH, France. He has published more than 220 research articles in peer-reviewed international journals and conferences.

Area of Interest/Research

Smart Grid, Artificial Intelligence techniques, Power Quality

Towards Quantum Artificial Intelligence

Electromagnetic Prediction Models for Ladder Logic Bombs and Faults in Programmable Logic Controllers

Varghese Mathew Vaidyan

Department of Electrical and
Computer Engineering, Iowa
State University, Ames, IA, USA
vaidyan@iastate.edu

Akhilesh Tyagi

Department of Electrical and
Computer Engineering, Iowa
State University, Ames, IA, USA
tyagi@iastate.edu

Abstract— Industrial control systems rely on Programmable Logic Controllers (PLCs) for the communication with sensors and actuators. On the other hand, vulnerabilities in Ladder logic and in turn PLCs being vulnerable to malware written in ladder logic or any other IEC 61131-3-compliant language is a new far more dangerous challenge in industrial control systems. Attackers install malware into a PLC's existing control logic to either modify the behavior continuously or wait for specific triggers. We consider Ladder Logic Bombs (LLBs) which are an extension of attacks like Stuxnet. These include stealthy LLBs, which are difficult to discover by humans manually evaluating PLC programs. We propose a monitoring technology for PLC's via Hybrid Quantum-Classical modelling of electromagnetic emissions from the Programmable Logic Controllers even in the case of stealthy LLBs. Our goal is to discover Ladder logic bombs from typical faults utilizing Electromagnetic (EM) spectrum domain-based control system processor fault analysis. We investigated various different types of ladder logic bomb activation, triggering modes, and payloads and discovered that our hybrid classical-quantum deep learning model can operate decoupled from the network and is very accurate at detecting logic layer vulnerabilities and faults.

Keywords— *deep learning, cyber-physical systems, industry applications, PLC, logic bombs, quantum computing, electromagnetics, fault analysis, control systems*

I. INTRODUCTION

Industrial Control Systems (ICS) that regulate power, production lines, manufacturing, and other vital infrastructure make use of Supervisory Control and Data Acquisition (SCADA) systems based on programmable logic controllers (PLCs) to communicate with sensors and actuators. Modern PLCs have sophisticated security features to ensure that only genuine firmware with verified signatures is deployed. However, ladder logic widely utilized in accordance with the IEC 61131-3 standard contains vulnerabilities since the logic implemented on PLCs may often be manipulated by an adversary with network/local USB access. This is a significant barrier for PLC security, even more so than adversarial attacks in traditional systems in IT, where it is more difficult to insert the attacker's code. Due to cyberwarfare, adversarial attacks and security problems in critical Cyber Physical Systems (CPS) are a prominent focus of interest presently [1, 2, 3–5]. The main

objective of this work is to detect a more potent attack on CPS employing PLCs in the form of logic layer attacks by ladder logic bombs (LLBs) using IEC 61131-3 ladder logic[6]. LLBs alter the behavior of a PLC or cause malicious activity to occur in response to a trigger signal. When the LLB is triggered, it substitutes altered values for actual sensor readings, causing the system to malfunction. The primary focus of this work is on attacks by stealthy LLBs (LLBs hard to detect by human operators who manually validate the ladder logic program).

An LLB based attacker may cause a Denial of Service (DoS) or modify the behavior of the PLC or obtain data traces of sensor and the control data executed by the PLC. This attack is more dangerous because it only requires a single point of entry in form of an intermittent (physical) access to the PLC. For instance, an attack can be made by a regular contractor having access to the PLC once a week as an example.

Along with LLBs, conventional ICS defects are a key source of contention. Without taking into account common defects, the system will suffer severe consequences. Model-based or parameter/state estimation-based techniques were used to identify ICS faults [7-9]. However, given the rising interconnection of the IoT, it is necessary to consider decoupled fault diagnosis approaches that are immune to adversarial attacks such as ladder logic bombs. Decoupled electromagnetic (EM) emission spectrum-based malware detection approaches have showed potential [10]. Our previous work has shown promise in control system processor fault identification [11-13]. But ladder logic bombs at the logic level is a recent far more dangerous stealthy attack in PLC and IEC 61131-3 ladder logic, after the Stuxnet and needs to extend the work to consider ladder logic bombs. The challenges arising from pipelining and also interaction by control system to external environment necessitates the need of Deep Learning frameworks to improve prediction accuracy.

Considering the capabilities of hybrid-Quantum deep learning frameworks, due to quantum entangling and angle embedding from quantum layer, hybrid quantum-classical has shown capabilities in predicting with higher accuracy [14]. Hybrid quantum-classical deep learning models of ladder logic bombs are proposed to overcome the challenges in detecting ladder logic bombs in PLC's. Experiments with the hybrid quantum fault models demonstrate an enhanced accuracy of

nearly 100% in differentiating ladder logic bombs from PLC failures.

II. RELATED WORK

Industrial Control Systems are used to monitor and control production lines, public infrastructure (water, electricity), and public transit systems. The programmable logic controller is a critical component of Industrial Control Systems. PLCs integrate sensor information and also execute control logic for the purpose of generating signals for actuators. It is crucial to protect PLCs against adversary attacks due to their widespread use and critical function in Industrial Control Systems. However, process control systems are susceptible to a variety of cyberattacks, some of which have the potential to result in catastrophic effects. [15, 16]. [17] examines vulnerabilities in SCADA control systems that use the MODBUS communication protocol. The majority of the attacks target measurement injection and denial-of-service attacks. Along with these risks, control logic vulnerabilities provide a significant threat, expressing as malicious logic alterations. Stuxnet [18] is a significant example of realistic control logic exploitation, creating a shift in the focus of advanced security measures. As a result, a renewed emphasis has been made on risks to PLCs based on control logic [19, 20, 21]. Karnouskos et al. analyzed Stuxnet and how it diverged from standard PLC behavior in [20]. The difficulties and impact of Stuxnet on nuclear power plants were highlighted in [21]. [19] discusses attacks that use probing requests to bypass authentication and obtain access to the PLC circuitry. Additionally, [22] investigates security risks to industrial PLCs at the network/firmware level. However, logic layer manipulations are becoming a major concern nowadays. Attacks based on PLC malware dynamically constructing a payload using the observations of the processes of the ICS is investigated in [23]. At first, the adversary malware collects needed details about the ICS process and also about the physical site. Following that, the attack is launched via the injection of a dynamic payload. However, in order to carry out this attack successfully, the attacker needs have prior information of the system being targeted. [24] develops an attack strategy based on automatically calculating the target PLC's needed semantics, which eliminates the need for advance knowledge of the system. [6] discusses the most sophisticated and stealthy ladder logic attacks similar to Stuxnet. [6] introduces the concept of ladder logic bombs using stealthy FFL and DoS attacks akin to Stuxnet.

Additionally, currently in place fault detection and identification approaches depend primarily on observer-based and canonical variate analysis-based approaches [25-28]. However, as IoT devices become more connected, a decoupled monitoring approach of faults and malware nodes becomes more capable of monitoring a large number of IoT devices owing to its ability to monitor the device from a distance.

Challenging classification problems have been solved using Quantum computing based deep learning approaches [29 - 31]. In this work, the classification and fault detection capabilities of hybrid QAI is utilized in classifying ladder logic bombs from traditional faults making the system suitable for critical infrastructure and applications.

III. HYBRID QUANTUM ARTIFICIAL INTELLIGENCE

The construction of a hybrid quantum classical quantum artificial intelligence model is represented in Figure 1. To convert conventional data into quantum states, the Data Embedding layer incorporates several qubit quantum gates including the Hadamard gate, Rotational X, Rotational Y, Rotational Z, as well as U1, U2, and U3 gates. In order to generate parameterized circuits, two-qubit gates (Controlled NOT, Controlled Z, and Controlled RX) are being used in combination with parameterized single-qubit gates. A quantum measurement component in the measurement layer establishes the qubit's current state. The outcomes of the quantum layer are then transmitted on to the subsequent layer as classical data [12].

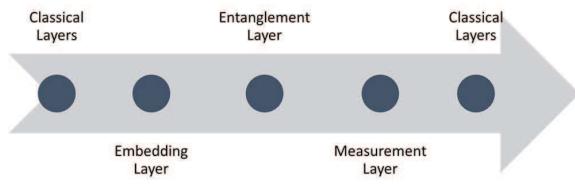


Fig. 1. Quantum Artificial Intelligence Process Flow

IV. LLB ATTACKS AND FAULTS

Digital signatures can make it more difficult to modify the firmware of Programmable Logic Controllers. However, a considerably more serious risk that is not adequately addressed by digital signatures is the security of the logic implemented on PLCs. Prior to uploading new logic to PLCs already in operation in an ICS, very little authentication are performed [6]. By exploiting this vulnerability, an attacker can launch an attack and effectively download bespoke logic onto the PLCs, jeopardizing the system and closing down vital infrastructure. Ladder Logic Bombs come in a variety of varieties [6]. The primary categories of LLBs considered in this paper are those that are triggered by system states, particular instructions or data, and the clock.

Payloads included into LLBs can be used to launch a Denial of Service (DoS) attack. To launch a Denial of Service (DoS) attack on PLCs, malicious logic is inserted into the ladder logic of the PLC and activated at a specific moment. Once launched based on an activation condition, the bomb enters an unending loop, rendering the PLC inoperable.

Another sort of LLB is used to covertly track and log critical PLC data. Utilizing FIFO buffers and storing data in the PLC as arrays is risky since they do not block system function and the operator is unaware of the LLB. These attacks are carried out in stealth mode and remain undetected for extended periods of time. However, they expose critical data and commands.

In terms of triggering, ladder logic bombs may be activated in response to the detection of a specified input. This paper

considers a scenario in which the device is triggered when a predefined value is attained. Another triggering circumstance that is examined is the detection of a certain trigger sequence. This exploit made use of Finite State Machine timers. It is difficult to detect these types of attacks since the exterior repercussions present themselves only after the damage has been done. Additionally, a TON timer-based LLB is triggered when the timer completes a count sequence. Another method of activating LLBs is through the use of a state variable that triggers the LLB when a certain state is reached.

Any time a controller bit remains at 0 or 1, it is referred to as a stuck-at fault. Normal functioning should include automatic shutoff and gradual reduction of levels to safe levels determined by sensors, for example. If the controller, on the other hand, is stuck at zero, it may perform an incorrect action, such as powering on the triggering unit and superheat to potentially lethal levels. For the industrial plant, this would be a disastrous incident, with the potential to be fatal. The focus of this research is to employ deep learning EM models of LLBs and stuck-at faults to detect potential stuck-at-0 or stuck-at-1 or LLB-caused failures in PLCs.

V. EM ANALYSIS TECHNIQUE

The process flow of the LLB detection in PLC's is shown in Figure 2. The PLC is executed continuously in no attack and LLB attacked modes for training trace data. The Electromagnetic emissions from the PLC are recorded using Electromagnetic probes. The acquired electromagnetic traces are converted to the frequency domain and fed into hybrid quantum deep learning systems. After that, the frequency-domain-based deep learning models are employed to construct models for LLBs. Following developing the models, they are used to interpret unknown PLC execution traces. The unknown traces are classed as LLB or fault related.

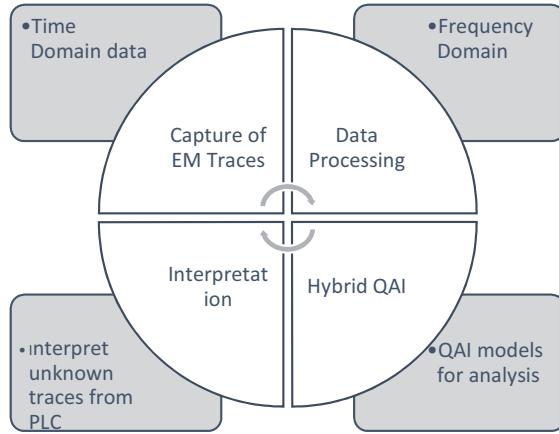
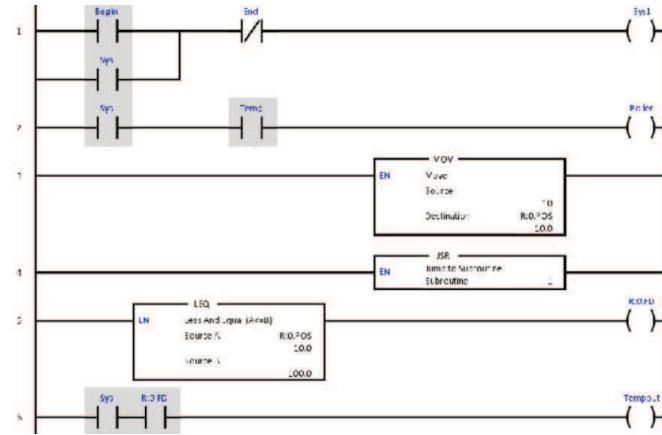


Fig. 2. Process Flow of PLC LLB Identification

VI. RESULTS AND DISCUSSION

Figure 3 shows the main routine of the boiler temperature control in PLC's. It works based on a setpoint value of temperature.

LLB_1: Stealth FFL logging is a type of LLB investigated in this study. It is a dangerous attack, and it is also more difficult to detect because there are no rapid changes in output. The adversary logs and exports data containing sensitive information in this type of assault. To replicate the attack, data recording was performed using an FFL-based FIFO buffer to store data in an array. The FFL block stores sensitive information about the process's setpoint value. An attacker with sporadic access can gain access to the values and use them to escalate the situation. A trigger condition based on a start contact as well as a trigger sequence were used to accomplish this. Because manual detection in large and sophisticated code with a high number of subroutines is difficult, the payload is also concealed within a subroutine in *LLB_1*.



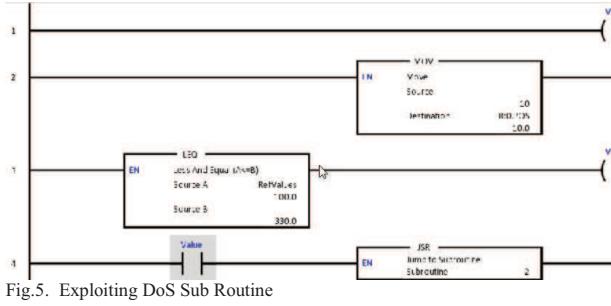


Fig.5. Exploiting DoS Sub Routine

LLB_2: A DoS attack was also simulated in order to disable a system's normal functionality. The system will be damaged if the PLC is unable to control the actuators. An indefinite loop was included as the bomb's payload to simulate a DoS attack. A specific data condition from the process was the triggering mechanism in this LLB. In this case, the LLB is hidden inside a function.

Stuck-at-Fault: To highlight the capability of the technique to distinguish LLB's from traditional faults, a Stuck-at-0 fault is evaluated. A stuck_at_0 was simulated by assuming the setpoint value is modified to a different value and the outcome of the execution of the ladder logic is altered.

The ladder logic was implemented on an Atmega microcontroller to test the technique's ability to discover vulnerabilities in ladder logic and PLCs. Because many PLCs use microcontroller-based architectures, the microcontrollers were chosen for examination. They were also used to highlight the ladder logic vulnerabilities. By fine-tuning the parameters of the neural network layers, the greatest prediction accuracy was reached. The optimal configuration (epochs=10, batch size=32, and verbose=1) resulted in the greatest accuracy. The optimal hybrid architecture consisted of four classical layers and a quantum layer comparable to [12].

The accuracy of the prediction in distinguishing a normal execution from *LLB_1* is shown in Figure 6. The Figure also illustrates the forecast accuracy for various triggering mechanisms. In all cases, hybrid quantum AI is extremely accurate at distinguishing *LLB_1* from regular execution.

The performance of the hybrid QAI models in detecting *LLB_2* from the regular operation of the ladder logic in a PLC under various triggering scenarios is shown in Figure 7. The hybrid QAI technique was capable of accurately distinguishing regular ladder logic execution in the PLC from an *LLB_2* DoS Attack. Table 1 examines the hybrid Q-AI model's fault/adversarial attack detection accuracy for multiple triggering methods, including *LLB_1* and stuck-at-fault. As shown in Table 1, the Hybrid model distinguished between *LLB_1* attack based on stealthy FFLs and a stuck-at-fault scenario with near 100 percent accuracy.

In Table 2, we examine the hybrid model's ability to distinguish between *LLB_2* (a denial-of-service attack) and stuck-at-fault under various triggering situations. Additionally, the Hybrid QAI model met the requirements in these cases.

These findings indicate that hybrid Q-AI model and technique performs well in the analysis of LLB's and faults in PLC's.

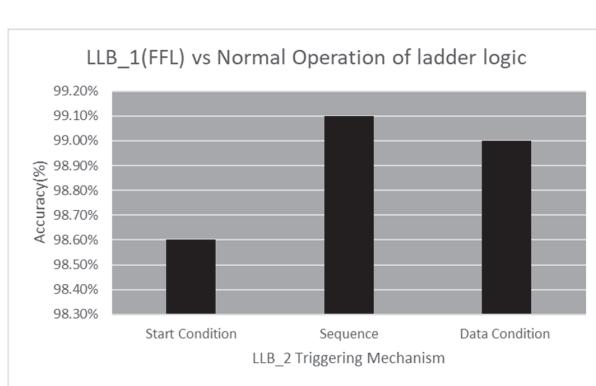


Fig. 6. Evaluation of hybrid Quantum-Classical models for distinguishing a normal PLC operation form an *LLB_1*

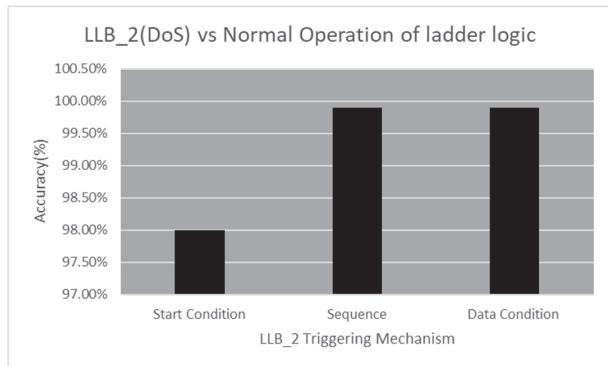


Fig. 7. Evaluation of hybrid Quantum-Classical models for distinguishing a normal PLC operation form an *LLB_2*

TABLE I. EVALUATION OF HYBRID QUANTUM-CLASSICAL MODELS IN DISTINGUISHING *LLB_1*(FFL) AND STUCK-AT-FAULT IN PLC

Triggering Mechanism	QAI Properties and Accuracy of Prediction	
	<Number of Qubits, Layers>	Accuracy of Prediction (%)
Data Condition	<4, 4>	99.2%
Sequence	<4, 4>	100%
Start Condition	<4, 4>	99%

TABLE II. EVALUATION OF HYBRID QUANTUM-CLASSICAL MODELS IN DISTINGUISHING *LLB_2*(DoS) AND STUCK-AT-FAULT IN PLC

Triggering Mechanism	QAI Properties and Accuracy of Prediction	
	<Number of Qubits, Layers>	Accuracy of Prediction (%)
Data Condition	<4, 4>	99.8%
Sequence	<4, 4>	100%
Start Condition	<4, 4>	99.7%

VII. CONCLUSION

In this paper, we examine ladder logic bomb vulnerabilities in ladder logic firmware, which are a significant problem in PLCs and are analogous to the alterations made by Stuxnet [18]. It was demonstrated in that even simple LLBs can be difficult to detect in real-world control logic code in [6]. Certain LLB attacks, such as the Stealthy FFL attack, are difficult to detect attributed to the reason that the outcome does not change immediately. However, the attack might continue undetected for years. Electromagnetic emissions reveal information about the integrity of a processor's computations. In classification problems, hybrid quantum classical deep learning models have been found to improve prediction accuracy. The capabilities of Hybrid Quantum Classical deep learning models in distinguishing between LLB and faults and LLB and normal operation of ladder logic was assessed. Due to the well-known theoretical advantages of quantum processing and the recent breakthrough in quantum hardware, quantum-enabled algorithms can increase the performance of classical machine learning models. The classical-quantum transfer learning model's performance was evaluated utilizing the controller's EM traces in Ladder logic bombs and failures. The analyses demonstrated that these models can maintain a high level of prediction accuracy when distinguishing Ladder Logic Bombs from other types of problems. This paves the path for novel methods of detecting stealthy ladder logic bombs such as Stuxnet.

In the future, more complicated attacks and faults on PLCs and RTUs can be assessed on the developed Hybrid models.

ACKNOWLEDGMENT

This work was supported in part by DARPA/AFRL through Contract FA8750-20-1-0504

REFERENCES

- [1] R. Chabukswar, B. Sinopoli, G. Karsai, A. Giani, H. Neema and A. Davis, "Simulation of network attacks on scada systems", Proc. 1st Workshop Secure Control Syst., pp. 587-592, 2010.
- [2] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On false data injection attacks against distributed energy routing in Smart Grid," 2012 IEEE/ACM Third International Conference on Cyber-Physical Systems, 2012.
- [3] E. K. Wang, Y. Ye, X. Xu, S. M. Yiu, L. C. Hui, and K. P. Chow, "Security issues and challenges for Cyber Physical System," 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing, 2010.
- [4] B. Zhu, A. Joseph, and S. Sastry, "A taxonomy of cyber attacks on SCADA systems," 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, 2011.
- [5] S. Zonouz, K. M. Rogers, R. Berthier, R. B. Bobba, W. H. Sanders, and T. J. Overbye, "SCPSE: Security-oriented Cyber-Physical State Estimation for power grid critical infrastructures," IEEE Transactions on Smart Grid, vol. 3, no. 4, pp. 1790–1799, 2012.
- [6] N. Govil, A. Agrawal, and N. O. Tippenhauer, "On ladder logic bombs in industrial control systems," Computer Security, pp. 110–126, 2017.
- [7] E. M. Cimpoeus, B. D. Ciubotaru, and D. Stefanou, "Fault detection and diagnosis using parameter estimation with recursive least squares," 2013 19th International Conference on Control Systems and Computer Science, 2013.
- [8] R. Isermann, "Model-based fault-detection and diagnosis – status and applications", Annual Reviews in Control, vol. 29, no. 1, pp. 71-85, 2005.
- [9] A. Mouzakitis, "Classification of Fault Diagnosis Methods for Control Systems", Measurement and Control, vol. 46, no. 10, pp. 303-308, 2013.
- [10] B. Yilmaz, R. Callan, M. Prvulovic and A. Zajic, "Capacity of the EM Covert/Side-Channel Created by the Execution of Instructions in a Processor", IEEE Transactions on Information Forensics and Security, vol. 13, no. 3, pp. 605-620, 2018.
- [11] V.M. Vaidyan, and A. Tyagi, "Instruction Level Disassembly through Electromagnetic Side-Chanel: Machine Learning Classification Approach with Reduced Combinatorial Complexity" In Proceedings of the 2020 3rd International Conference on Signal Processing and Machine Learning, 2020 .
- [12] V.M.Vaidyan and A. Tyagi, "Hybrid Classical-Quantum Artificial Intelligence Models for Electromagnetic Control System Processor Fault Analysis", IEEE IAS Global Conference on Emerging Technologies, In press, 2022.
- [13] V.M.Vaidyan and A. Tyagi, "Failure Analysis of Control System Processors", International Journal of Cybernetics and Cyberphysical Systems, In press, 2022.
- [14] Quantum Artificial Intelligence in 2021: in-Depth Guide. 2020 [cited 2021 Aug 1]. Available from: <https://research.aimultiple.com/quantum-ai/>
- [15] S. Amin, X. Litrico, S. S. Sastry and A. M. Bayen, "Stealthy deception attacks on water SCADA systems: Proceedings of the 13th ACM International Conference on Hybrid Systems: Computation and control," ACM Conferences, 01-Apr-2010.
- [16] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," ACM Transactions on Information and System Security, vol. 14, no. 1, pp. 1-33, 2011.
- [17] T. H. Morris and W. Gao, "Industrial Control System Cyber attacks," Electronic Workshops in Computing, 2013.
- [18] N. Falliere, L. O. Murchu, and E. Chien. W32. stuxnet dossier.
- [19] D. Beresford. Exploiting Siemens Simatic S7 PLCs, 2011. Proceedings of Black Hat USA.
- [20] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, 2011.
- [21] D.-Y. Kim, "Cyber security issues imposed on nuclear power plants," Annals of Nuclear Energy, vol. 65, pp. 141–143, 2014.
- [22] S. A. Milinkovic and L. R. Lazic, "Industrial plc security issues," 2012 20th Telecommunications Forum (TELFOR), 2012.
- [23] S. McLaughlin, "On dynamic malware payloads aimed at programmable logic controllers", in proceedings of 6th Workshop on Hot Topics in Security, USENIX, San Fransisco, USA, Aug. 2011
- [24] S. McLaughlin and P. McDaniel. SABOT: Specification-based payload generation for programmable logic controllers. In Proc. of the ACM Conference on Computer and Communications Security (CCS), pages 439–449. ACM, 2012.
- [25] I. Hwang, S. Kim, Y. Kim and C. E. Seah, "A Survey of Fault Detection, Isolation, and Reconfiguration Methods," in IEEE Transactions on Control Systems Technology, vol. 18, no. 3, pp. 636-653, May 2010
- [26] P. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy—A survey and some new results", Automatica, vol. 26, no. 3, pp. 459-474, May 1990.
- [27] E. Bernardi and E. J. Adam, "Observer-based Fault Detection and Diagnosis Strategy for Industrial Processes", Journal of the Franklin Institute, vol. 357, no. 14, pp. 9895-9922, 2020.
- [28] L. Li, S. X. Ding, J. Qiu, K. Peng and Y. Yang, "An optimal fault detection approach for piecewise affine systems via diagnostic observers", Automatica, vol. 85, pp. 256-263, Nov. 2017.
- [29] Y. Dang, N. Jiang, H. Hu, Z. Ji and W. Zhang, "Image classification based on quantum K-Nearest-Neighbor algorithm", Quantum Inf. Process., vol. 17, no. 9, pp. 239, Sep. 2018.
- [30] Z. Khan, S.M. Khan, J.M Tine, A. T. Comert, D. Rice, G. Comert, D. Michalaka, J. Mwakalunge, R. Majumdar, and M. Chowdhury, "Hybrid Quantum-Classical Neural Network for Incident Detection" arXiv preprint arXiv:2108.01127, 2021.

- [31] J. Amin, M. Sharif, N. Gul, S. Kadry and C. Chakraborty, "Quantum machine learning architecture for COVID-19 classification based on synthetic data generation using conditional adversarial neural network", Cognit. Comput., vol. 14, pp. 1-12, Aug. 2020.

Deep Learning Approaches for Crack Detection in Bridge Concrete Structures

Daniel Einarson

Department of Computer Science
Kristianstad University
Kristianstad, Sweden
daniel.einarson@hkr.se

Dawit Mengistu

Department of Computer Science
Kristianstad University
Kristianstad, Sweden
dawit.mengistu@hkr.se

Abstract—Convolutional Neural Networks are among the most effective algorithms for image analysis applications. However, the accuracy of the algorithms depends on the availability of powerful computational resources and the quality of the images used to train the models. This paper investigates ways to build robust models to detect cracks in concrete structures using low resolution images and third-party datasets. Our experiments show that reducing image sizes by a factor of 4 does not significantly impact the accuracy. This is helpful to shorten execution time and hence lower cloud service costs. It is also observed that applying a model trained on one image dataset to detect cracks in images from a different source is not a trivial task.

Keywords—Concrete Crack Detection; Convolutional Neural Networks; Deep Learning; Performance Optimization

I. INTRODUCTION

Convolutional Neural Networks (CNN) have been successfully used for image analysis in different application domains. The use of CNN models to detect cracks in concrete structures has been investigated and promising results have been achieved [1]. These models can be effectively employed to detect existing cracks and predict potential damages in roads railways, buildings, bridges, etc. [2], [3]. Machine learning models have a major advantage over the traditional manual approach used for crack detection. Manual inspection compromises safety as it exposes people to work in dangerous and inaccessible parts of the concrete structures. Automating the process by employing robots to capture images in hazardous and inaccessible areas and feeding these images to a machine learning pipeline is recommended to address this problem.

This paper presents a study that investigates ways to improve the accuracy of concrete crack detection on the Öresundsbron (the Öresund Bridge). The Bridge management is tasked with performing inspection of the bridge for preventive maintenance. The inspection involves early detection of cracks in the bridge's structures. It is intended to automate this task by using drones to take images of the concrete structure, and by applying machine learning techniques on these images for crack detection.

The objective of this study is to investigate how machine learning models can be implemented and incorporated into the Bridge's preventive maintenance system. From a research

perspective, the main issues investigated in this work are classification accuracy of the model and, improvements of computational performance, mainly in the cloud. The rest of this paper is organized as follows. A brief background on the Öresund Bridge and review of CNN models is presented. This is followed by descriptions of the experimental methodology applied in this study. The achieved results are then presented along with discussions on their significance. The report concludes by summarizing the important findings and citing directions for future work.

II. BACKGROUND

A. Brief Overview of the Öresund Bridge

The Öresund Bridge is a combined railway and motorway bridge that links Sweden and Denmark. The bridge runs nearly 8 kilometers (5 miles) from the Swedish coast to the man-made island Peberholm. The island is connected to the Danish coast via the 4-kilometre (2.5 mi) Drogden Tunnel on its other end. Öresundsbro Konsortiet¹ is the management body responsible for maintaining the bridge, and related assets.

At present, the bridge management has a large repository of footages that were not systematically organized. Because the images were taken from afar, they lack clarity and sharpness desired for an accurate analysis. As a result, it has been difficult to draw meaningful conclusions using the repository in its present form.

A pilot study to investigate the possibilities of using CNN on this dataset was performed as a thesis work at Kristianstad University [5]. A dataset of 6,639 images (4,633 crack free and 2,006 having cracks) was available for this study. The original images in the dataset have a 6013x3376 pixels size. However, they were cropped down to 256x256 pixels for ease of analysis, inspired by the work of [1]. Fig. 1 shows few examples of images containing cracks vs crack-free.

The study showed that there are two major challenges to achieve the intended objective. First, because the study was performed on a laptop with moderate resources (CPU, memory), the model training phase took a long time. Second, it was not possible to achieve the desired level of accuracy owing to the limited size of the training dataset. The quality of the images and

¹ Öresundsbro Konsortiet - <https://www.oresundsbron.com/en/info/company>

the balance between image classes (number of with-crack vs. no-crack) also weighed on the accuracy of the models.

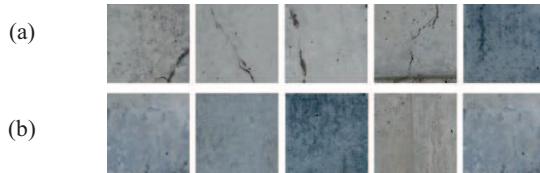


Fig. 1. Images from Öresund dataset. (a) with Cracks. (b). cracks-free

As is the case in image classification projects, tackling these limitations is of significant interest. Deploying the model on the Cloud has become the preferred approach as the Cloud proved to be an efficient high performance computing environment. To address the accuracy aspect, we build the CNN model using a third-party dataset and test the model's performance with the images from our Bridge dataset. Accordingly, we shall use the Mendeley dataset², which is available in a public repository and used for similar studies [4]. Fig. 2 illustrates representative images from the Mendeley dataset to show the difference between with-crack and no-crack image classes.

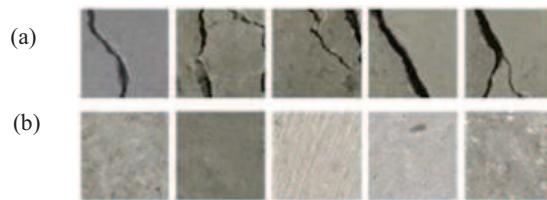


Fig. 2. Images from Mendeley dataset. (a) with Cracks. (b). cracks-free

B. Review of Convolutional Networks

While CNNs can be applied in several application domains, they have shown superior performance in image recognition, due to their built-in structures [6]. A CNN basically consists of Convolutions, Pooling, Activation Functions, and Dense Layers. These operations are performed in a number of layers through which the original input data is matched against possible output, as illustrated by Fig. 3. Such layers are often supplemented with a final *Softmax* operation before mapping to Output. For more information about this, see, [1] and [7].

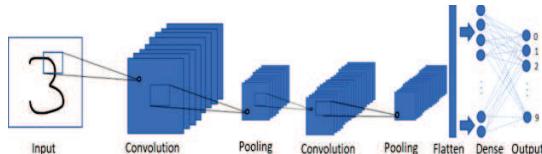


Fig. 3. A CNN to classify handwritten digits

Fig. 3 relates to the classic MNIST-example [8], which examines neural networks to recognize hand-written digits of size 28x28 pixels. A study in [8] tackles the problem on the basis of variations of the LeNet CNN-structure (Fig. 3). Although a LeNet-based CNN may reach an accuracy above 99%, the

MNIST-problem may actually be solved by a significantly simpler and faster, fully connected neural network. A simple feed forward network with only one hidden layer achieves an accuracy level of about 98% for this problem [9].

Techniques relying on simple networks to detect cracks were proposed by [11]. However, this is not the case for our crack detection dataset. This is because the images from the Öresund Bridge are significantly larger and too blurry to be handled by simple networks. Therefore, the need for advanced algorithms such as convolutional neural networks is evident.

Convolutional networks are becoming increasingly complex and evolved with the growing challenges in image classification problems [12]. Earlier versions of CNN such as LeNet-5 (1998), AlexNet 2012 and VGG16 (2014) consist of a sequence of a lower number of layers. Later variants of CNNs such as Inception-v3 (2015), and ResNeXt-50 (2017) have complex structures consisting of combinations of layers organized in parallel.

Previous studies on crack detection show that CNNs of the type VGG-16 can produce accurate models [1]. It can be observed in the results in [13] that VGG-16 achieved a good accuracy on a dataset of 2500 images with 256x256 resolution. The CNN in this study has 13 convolutional, and 3 fully connected layers. An accuracy of 92.27% is achieved in 50 epochs. An AlexNet network consisting of 5 convolutional, and 3 fully connected layers proposed in [14] achieves an accuracy level of about 98%. Surprisingly, the more advanced VGG16 models actually show lower accuracy than the less advanced AlexNet.

The study in [1] gives interesting insights about the accuracy and execution performance of CNNs. The model developed in this study has 8 layers: 4 convolutional, 2 pooling, 1 ReLU, and 1 Softmax layers. The concrete images were taken at a close distance (1 – 1.5 meters from the concrete). In total, 32K images were used, achieving an accuracy of about 98% around the 50th epoch. The experiment completed in 90 minutes on two GPUs while it took 1-2 days on a standard CPU. This shows the need for accelerated computing environment, to achieve speed and performance.

A study in [4] compares several types of CNNs, such as AlexNet, VGG16, and ResNet50, based on the Mendeley image dataset, and shows impressive results, with an accuracy of above 99%. The dataset contains 40K high quality images, with 227x227 pixel resolution. The time required to train 28K dataset per epoch is, however, unacceptably high. For instance, the time required for AlexNet is shown to be 133 seconds, and for VGG16, it is 2,827 seconds.

From the above studies, it can be seen that a simpler CNN such as AlexNet can achieve an acceptable accuracy and performance. However, early experiments with the Öresund Bridge images of size 256x256 pixels showed unsatisfactory results. Worse to note that the time to train the system to get such dismal results was about 24 hours (on one machine), pointing to the clear need for alternative approaches to address the problem.

² Mendeley Data set for Classification of Concrete Crack Images, <https://data.mendeley.com/datasets/5y9wdsg2zt/2>

III. METHODOLOGY

We conducted experiments in two different computational environments, to investigate execution performance issues. The first experiment is performed on a laptop computer (we call it the local version) having 11th Generation Intel core i5-11400H, 2.30GHz processor, 12MB cache and 16 GB RAM.

The second experiment is performed in the Amazon cloud on custom-built, accelerated computing instance of *ml.g4dn.4xlarge* type. This powerful machine has 16 virtual CPUs, 64GB RAM and a T4 GPU (2560 cores). This cloud instance is provisioned with AWS SageMaker Notebook preconfigured for deep learning in the Python language. In both cases, our training models are based on an AlexNet-type CNN consisting of six layers as follows:

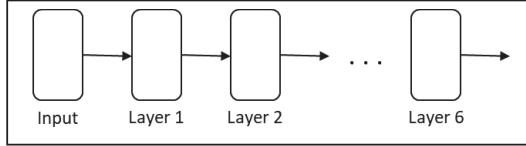


Fig. 4. A six-layer AlexNet-type of CNN

1. A Convolution layer of size 16 (C-16), an Activation function (A) of type Relu, a MaxPooling (P) of size 5x5, and a DropOut (D) of 20%
2. C-32, A-Relu, P-3x3, D-20%
3. C-64, A-Relu, P-2x2, D-20%
4. A Flatten Layer
5. A Dense Layer of size 32, and A-Relu
6. Dense-1, and a Sigmoid Activation Function finally providing output.

The experiments were run for 20 epochs, with learning rate set to 0.001 for the first 10 epochs and decreased to 0.0005 for the rest³. The batch size chosen is 64.

The original Mendeley dataset⁴ contains 40 000 images split into 20 000 Positive (with cracks), and 20 000 Negative (without cracks). We selected some 10K images from this dataset to conduct our experiment. The following two-step approach was further employed.

1. To investigate the effect of image size on classification accuracy and execution performance, the original images were resized to different sizes and used to train the CNN (procedure A).
2. For each category of resized images, the experiment was repeated by varying the proportion of positive images (containing crack) used to train the model. This step was used to investigate the effect of unbalanced data on classification accuracy (procedure B).

Both procedures were implemented and evaluated on both the local and cloud-based experiments.

³ A learning rate of 0.001 is a standard setting, tuning the value is a common way to further improve accuracy.

A. Resizing the images

As explained earlier, we chose 5100 positive (having crack) images as well as 5100 negative (crack-free) images from the Mendeley dataset. Three test-cases are prepared with 100 images removed from each group and saved aside for use as test images. Therefore, we have three different sets of 5000 (positive plus negative) images for training, and an additional 100 images (positive plus negative) for testing. From the training set, 20% of the images were used for validation, thus, the other 80% are used for building the CNN model.

Next, we created new datasets of lower resolution from the original 227x227 images. The new image sets have sizes of 128x128, 64x64, and 32x32 pixels. Accordingly, a test-suite of twelve different cases is created, that is, four image-sizes times three test-cases for each. The reference scale on which the images are resized is the original size of the Öresund Bridge images, which is 256x256 pixels. Halving the sizes of the images is considered significant enough to observe clear effects. The resizing is done through an *INTER_AREA* method, as described in [16].

The outcome of image resizing is illustrated in Fig. 5. A crack is seen on the original and the resized images. The assumption here is that, if cracks are clearly visible to the eye, they should certainly be detected by a CNN model as well.

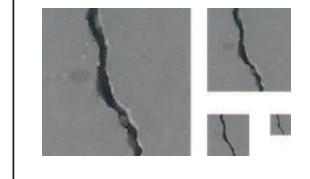


Fig. 5. Different size images with cracks

B. Unbalancing the datasets

In the real world, cracks in concrete are a seldom occurrence. This means, the number of images with cracks is certainly much less than that without cracks if the dataset is assembled naturally. This imbalance poses clear difficulties in training the CNN model in the usual way. The use of unbalanced classes of data (positive vs. negative) has obvious impact on accuracy.

In this part of the experiment, we start with completely balanced sets of images 5000 positive plus 5000 negative and proceed to lower the number of positive images in the next runs. This is achieved by successively removing positive images so that the number of positive images is 40%, 20%, 10%, 5%, 1% and 0.1% of the negative images and then rerun the test. For instance, an imbalance factor of 40% corresponds to 5000 negative, and 2000 positive images.

The accuracy figures should be regarded with special care. For instance, in a case where the number of positive images in the dataset is less than 10% of the negative images, an accuracy of 90% may not be satisfactory. Such an accuracy result could

⁴ The Bridge dataset has 4,633 images considered without cracks, resp. 2,006 images considered with cracks.

lead to reporting many false negative or false positive cases. Therefore, the result must be complemented with the confusion matrix⁵.

IV. RESULTS

Our results are reported below according to the structure outlined in the Methodology in Section III. The tables show size of images, time to train and validate a CNN in seconds, percentage of training accuracy (*Accuracy*), and *Validation accuracy*, for all test cases 1-3. The details from the confusion matrix tables (*TP*, *TN*, *FP*, *FN*) are presented for both procedures (experiments with resized images and unbalanced datasets).

A. Local (laptop-based) experiments

TABLE I, shows the experiments run for the three testcases, and for the different image sizes. It can be observed that the values of the training as well as the validation accuracy are surprisingly high. That is, due to the sharpness of the images of the Mendeley dataset, even downsampled images give high accuracy.

As can be seen, the execution time and image size are not linearly related. Doubling image size by a factor of about 2, implies quadrupling the number of input nodes and hence the total number of edges will be accordingly higher. This seems to result in a difference in execution time by a factor of about 3. This is clearly demonstrated in TABLE I.

TABLE I. RESIZING IMAGES THROUGH DOWNSCALING

<i>Case</i>	<i>Size</i>	<i>Exec. Time (sec)</i>	<i>Accuracy (%)</i>	<i>Val. Accuracy (%)</i>
1	32	72.3	98.4	98.7
1	64	225.9	99.2	98.8
1	128	819.3	99.7	99.0
1	227	2556.2	99.8	99.3
2	32	77.9	98.5	99.0
2	64	257.7	99.3	98.9
2	128	844.0	99.4	98.3
2	227	2551.9	99.6	99.0
3	32	79.7	98.6	99.1
3	64	261.9	99.2	98.9
3	128	844.0	99.5	98.9
3	227	2567.7	99.7	98.9

The accuracy results need to be interpreted to evaluate their significance. Therefore, we generated the data in TABLE II so that conclusions can be drawn as to whether the obtained results are fair enough for all image sizes.

TABLE II. CONFUSION MATRIX FOR RESIZED IMAGES

<i>Case</i>	<i>Size</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>
1	32	99	96	4	1
1	64	100	98	2	0
1	128	100	99	1	0
1	227	100	98	2	0
2	32	91	100	0	9
2	64	97	100	0	3
2	128	100	99	1	0
2	227	99	99	1	1
3	32	100	97	3	0
3	64	100	97	3	0
3	128	100	98	2	0
3	227	99	100	0	1

TP=True Positive, *TN*=True Negative,
FP=False Positive, *FN*=False Negative

For procedure B, image sizes of 64x64 pixels are used to run the experiments with unbalanced images sets. The results of these experiments for different levels of imbalance are shown in TABLE III. It is interesting to note the achieved high accuracy values and, low false positive and false negative rates in these experiments.

From TABLE III, it can be observed that the images of the Mendeley dataset are sharp enough to give high accuracy with unbalanced datasets. The false negative rate is low if the imbalance factor remains above 10%. It is important to note that the accuracy figure alone can be deceptive. As can be seen in the table, a test with an imbalance factor of 0.1% gives 99.9% accuracy, while its false negative rate is 100 (or predicted all crack images wrong).

TABLE III. UNBALANCING SETS OF IMAGES OF SIZE 64X64 PIXELS

<i>Imbalance factor (%)</i>	<i>Accuracy (%)</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>Exec. Time (sec.)</i>
40	99.5	100	96	4	1	154.5
20	99.5	99	99	1	1	134.2
10	99.4	98	100	0	2	115.5
5	99.7	96	99	1	5	114.1
1	99.6	31	100	0	69	114.2
0.1	99.9	0	100	0	100	110.5

Furthermore, the values *TP*, *TN*, *FP*, and *FN*, may be used to identify further measurements, such as *Sensitivity (Sv)*, and *Specificity (Sp)*. Those are defined as follows ([15]):

$$Sv = TP / (TP + FN) \quad (1)$$

⁵ Please, see, e.g., [15] for more information on the concepts of the confusion matrix.

$$Sp = TN / (FP + TN) \quad (2)$$

In the context of the results shown in TABLE III, Sv explains how accurately images showing cracks can be identified by our model. Conversely, Sp refers to the model's ability to correctly identify crack free images. Table IV shows these values computed from the results shown in TABLE III.

TABLE IV. EXTENSION TO TABLE III

Imbalance factor (%)	Sensitivity	Specificity
40	99%	96%
20	99%	99%
10	98%	100%
5	95%	99%
1	31%	100%
0.1	0%	100%

It can be seen in TABLE IV that the sensitivity of the model declines drastically as the proportion of positive images (imbalance factor) in the training set is reduced. This is expected because the network is exposed mostly to negative images during the training phase. On the other hand, the model is robust for crack-free images.

B. Cloud based experiments

In this part of our work, the laptop-based experiment is repeated on the cloud albeit with different sizes of data. The aim of this task is to obtain a more accurate model by using a much larger dataset. Additionally, it is interesting to observe if the model trained on images from one concrete structure can be used for prediction in a different structure. This possibility would simplify the job of practitioners who have a limited dataset to train a new model of their own. It is also beneficial for those users whose image dataset is not yet labelled as manual labelling of a large dataset requires opening and visually inspecting each image file. Accordingly, we used images from two different sources for the test set. The first test case used the Mendeley dataset, similar to the laptop- based experiment. The second test case used images from the Bridge dataset itself. The experiment was repeated for different image sizes. The results of both test cases are shown in figures 6 and 7.

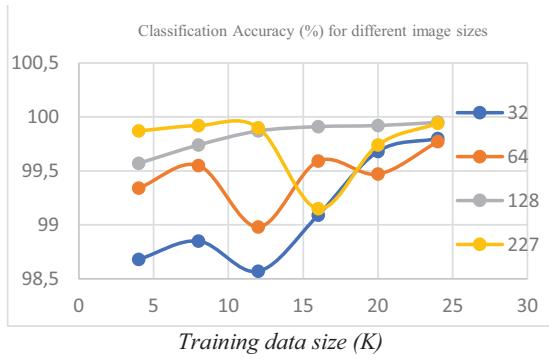


Fig 6. Experiment with Mendeley images

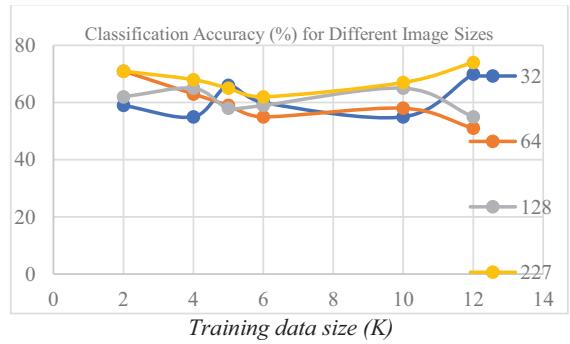


Fig 7. Experiment with Bridge images

The trained model gives a good prediction accuracy (close to 99%) when tested with Mendeley images. However, the accuracy of the model on the actual Bridge images is not satisfactory. A closer examination of the images shows that the Bridge images were blurry and therefore, the training dataset needs further processing to accommodate this fact.

Another important observation we made is that the model has an accuracy above 90% when tested with the crack-free (negative) images from the Bridge dataset itself. However, the accuracy is very low for crack containing (positive) images of the Bridge. This is an indication of a high rate of false negatives and thus our model cannot reliably detect cracks. We explored other options such as increasing the proportion of positive images, to train the model with more images showing cracks. However, the accuracy improvement is not significant.



Fig 8. Execution performance on cloud

An important observation we made on the cloud execution environment relates to performance. As expected, the cloud-based implementation is much faster than the lap-top version. It could train 24K images in less than 12 minutes on the 227X227 set. This is in fact many folds improvement over the local version. It is also possible to see that the experiment with 32X32 images executes 10 times faster than the 227X227 set while there is no significant difference in their accuracies. Achieving comparable accuracies with low resolution images has a significant impact as there will be substantial reduction in cloud services cost.

Another important observation we made is that transferring large datasets from the repository to the cloud execution node could take longer time than the actual training time. This incurs significant costs specially on repeated and long-running experiments. Furthermore, due to the inherent inefficiency of Python's memory management, vectorizing the images to prepare them for training consumes a substantial part of the memory. We were able to tackle these challenges and managed to train large datasets by implementing improved data caching and memory optimization techniques.

V. CONCLUSIONS AND FUTURE WORK

This study addressed two major issues in machine learning: accuracy, and performance of CNN models in concrete crack detection use case. Because the dataset provided by the user was not in a readily usable form, the Mendeley crack dataset was used in the study. The experiment was tested on a laptop for proof-of-concept and a larger version was deployed and evaluated on the cloud. Resizing the images did not show significant loss of accuracy. The 64X64 images give an accuracy level of 99.2%, while their execution time is less than 10% of the 227X227 images.

It has been shown that the original images of 227x227 pixels can be scaled down to 64x64 pixels without any significant loss in accuracy, and with extraordinary results (model accuracy about 99.2%), compared with previously published studies. Furthermore, the effect of unbalanced datasets has been studied by progressively decreasing the proportion of positive images in the training set. It was found that the prediction accuracy is robust even when the model is trained on a dataset with only 10% positive images. Preliminary studies on the Öresund Bridge dataset also show promising results, with about 98% accuracy, on images scaled down to 64x64 pixels, and an imbalance factor in the training dataset of about 40%.

While this work can be considered a preliminary study, it yielded important results and the experimental findings opened additional research inquiries for future work. Relevant questions to ask in this regard include: *How far can one simplify the CNN model without compromising accuracy? How can the relation between image-sizes and CNN-structure on one side, and execution time on the other be balanced? How can we build an accurate prediction model on one dataset and use it on images from other sources than it trained on? How can we reduce the effect of dataset imbalance for this particular use case?*

ACKNOWLEDGMENT

The authors would like to acknowledge students involved in the initial experiments of this project. Special thanks also to the managers of the Öresund Bridge for interesting and valuable discussions and support.

REFERENCES

- [1] Cha Y, Choi W, Büyüköztürk O. "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks". Computer-Aided Civil and Infrastructure Engineering. 2017;32(5):361-378.
- [2] Lei Zhang , Fan Yang , Yimin Daniel Zhang, and Y. J. Z., Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). "Road crack detection using deep convolutional neural network," 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 3708-3712, <http://doi.org/10.1109/ICIP.2016.7533052>
- [3] Soukup, D. & Huber-Mork, R. (2014), Convolutional neural networks for steel surface defect detection from photometric stereo images, in Proceedings of 10th International Symposium on Visual Computing, Las Vegas, NV, 668-77.
- [4] Özgenel, Ç.F., Gönenç Sorguç, A. "Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings", ISARC 2018, Berlin.
- [5] Martijn de Redelikheid Kristian Kokoneshi, A Machine Learning Analysis of Photographs of the Öresund Bridge, Bachelor Thesis at Dept. of Computer Science, Kristianstad University, 2020, available at <http://hkr.diva-portal.org/smash/record.jsf?pid=diva2%3A1451429&dswid=-6276>
- [6] LeCun, Yann & Bengio, Yoshua, Convolutional Networks for Images, Speech, and Time Series, The Handbook Brain Theory Neural Networks, vol. 3361, 1995.
- [7] Saha Sumit, A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way, towards data science, Dec 15, 2018, available at: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- [8] LeCun, Yann; Léon Bottou; Yoshua Bengio; Patrick Haffner (1998). "Gradient-based learning applied to document recognition" (PDF). Proceedings of the IEEE. 86 (11): 2278–2324.
- [9] Rashid T., Make Your Own Neural Network, Createspace Independent Publishing Platform, ISBN 9781530826605, 2016.
- [10] Neural Network, Databricks, No date, available at <https://databricks.com/glossary/neural-network>
- [11] Hoang N-D., Detection of Surface Crack in Building Structures Using Image Processing Technique with an Improved Otsu Method for Image Thresholding, Advances in Civil Engineering, Volume 2018 | Article ID 3924120 | <https://doi.org/10.1155/2018/3924120>, 2018.
- [12] Karim Raimi, Illustrated: 10 CNN Architectures - A compiled visualization of the common convolutional neural networks, towards data science, Jul 29, 2019, available at: <https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614d>
- [13] Silva W, Lucena D. "Concrete Cracks Detection Based on Deep Learning Image Classification." Proceedings. 2018;2(8):489
- [14] Kim H, Ahn E, Shin M & Sim S. "Crack and Noncrack Classification from Concrete Surface Images Using Machine Learning." Structural Health Monitoring. 2018;18(3):725-738.
- [15] Géron A., Hands-On Machine Lerning with Scikit-Learn, Keras & TensorFlow. O'Reilly, ISBN 978-1-492-03264-9, 2019.
- [16] Dong W., What is OpenCV's INTER_AREA Actually Doing?, Medium, 2018, available at: <https://medium.com/@wenrudong/what-is-opencv-inter-area-actually-doing-282a626a09b3>

On-Demand Recharge Scheduling Algorithm in Wireless Sensor Networks

Bhargavi Dande*

Corresponding author

Dept. of Computer Science
and Information Engineering
Tamkang University
New Taipei City, Taiwan
808415011@gms.tku.edu.tw

Chih-Yung Chang
Dept. of Computer Science
and Information Engineering
Tamkang University
New Taipei City, Taiwan
cychang@mail.tku.edu.tw

Ming-Yang Su
Dept. of Computer Science
and Information Engineering
Ming Chuan University
Taoyuan, Taiwan
minysu@mail.mcu.edu.tw

Cheng-De Fan
Dept. of Electrical and
Computer Engineering
Tamkang University
New Taipei City, Taiwan
408500477@gms.tku.edu.tw

Abstract— In recent years, energy transfer technology provides a new opportunity for recharging the sensors using the mobile charger. Most of the studies proposed recharging algorithms using a single charger but overlooked the consideration of multiple chargers and cooperation between them. Therefore, this paper proposed an On-Demand Recharge Scheduling (ODRS) algorithm for multiple mobile chargers, aiming to improve the network surveillance quality (NSQ). In this paper, the network is partitioned into many subregions. One mobile charger is employed in each subregion to recharge the sensors based on the benefit of coverage quality (BCQ) obtained. In addition, there is cooperation between neighboring mobile chargers. Experiments are evaluated to show the achieved NSQ of the proposed ODRS algorithm.

Keywords— *on-demand recharging, wireless sensor networks, multiple mobile chargers, routing load, coverage quality.*

I. INTRODUCTION

Wireless sensor networks (WSNs) were utilized in many applications like smart homes, industries, agriculture, etc. Although they were used in a lot of applications, the restricted energy of sensors degrades the WSNs. Therefore, how to recharge the sensors nodes is one of the major issues, which needs to be further investigated. In the literature, the sensors were recharged by the mobile chargers using energy transfer technology. The literature of these studies is categorized into two classes: periodic and on-demand charging schemes. In the periodic schemes [1] – [5], the mobile chargers travel to all the sensors and recharge them periodically. These types of studies ignored that each sensor has distinct energy consumption. In addition, the mobile charger consumes a lot of energy for its movement.

To overcome the issues faced by the periodic charging schemes, on-demand charging schemes [6] – [12] were proposed. These studies define the threshold value for sensor batteries to forward the request. Therefore, the mobile charger travels only to the requested sensors. That is to say, these studies considered real-time charging requests. In the on-demand charging schemes, there might be many requests received to the mobile charger, especially in large-scale WSNs. Therefore, managing the recharging schedule and selecting the sensors, which need to be recharged is the most important issue.

II. RELATED WORKS

The literature on energy transferring technologies was divided into periodic and on-demand recharging schemes.

A. Periodic Charging Schemes

Study [1] proposed a multi-node recharging algorithm to

reduce the dead nodes and energy usage. This study assumed that the sensors were partially recharged and all the sensors in the charging range could be recharged. In addition, this study allows the mobile charger to recharge the sensor multiple times. Study [2] investigated to find the stop positions and path planning of the mobile charger. Another study [3] proposed an algorithm to design the trajectory for chargers. They considered the routing load (RL) of the sensors to recharge them. This study stated that the sensors closer to the base station (BS) have higher RL, which leads to consuming more energy. Therefore, the path length of the mobile charger assigned to the sensors closer to the BS should be very short to recharge all of them timely.

Study [4] proposed an algorithm to improve the network lifetime. This study proposed two algorithms, the first algorithm considered the remaining lifetime of sensors. The second algorithm focus on reducing the charging time. Another study [5] determines the rendezvous points as the stop locations. All the sensors in the charging range of the rendezvous points could be recharged. However, all the above studies assumed that the energy discharging rate is constant for all the sensors, which is not very practical.

B. On-Demand Charging Schemes

Study [6] proposed an on-demand recharging algorithm based on the nearest job next. That is, the sensor nodes closer to the mobile charger location will be recharged first to reduce the path length. Study [7] proposed a coverage-aware algorithm for WSNs. This study partitions the network into equal-sized grids and assigns weight to each grid. The mobile charger travels to the grids based on the maximum weight and obtained coverage benefit.

Study [8] adopted fuzzy logic to investigate the on-demand recharging schedule. This study considered parameters such as energy, location of the mobile charger and node density to schedule the mobile charger. Another study [9] proposed a multi-mobile chargers scheduling algorithm with time windows. This paper assumed that the mobile chargers have limited battery capacity. Thus, aimed to minimize the moving cost and time windows violation. Study [10] dynamically reconstructs the path by inserting and deleting the sensors based on the coverage of sensors. Although studies [6]-[10] reconstruct the mobile charger path dynamically, they fixed the threshold value of sensors, regardless of the real-time information of the sensors.

Study [11] designed two threshold values for each sensor. Based on the threshold values, this study assigned the priorities of sensors to be recharged. In addition, this study

used real-time information such as the length of the queue and energy discharging rate to design the threshold values. However, they ignored the cooperation between mobile chargers. Another study [12] focused on multiple charger scheduling algorithms for different size clusters. The clusters were divided into many levels according to the remaining energy of the sensors. This study determines the chargers according to the requests received.

The proposed on-demand recharge scheduling (ODRS) algorithm firstly partition the network based on the RL of the sensors. Then adopted the queuing theory to determine the threshold value of sensors. Finally, the mobile chargers construct their paths based on the benefit of coverage quality (BCQ) obtained from the sensors. In addition, there is cooperation between neighboring mobile chargers.

III. NETWORK ENVIRONMENT

This study assumes WSNs, where the area is denoted by W . Let $S = \{s_1, s_2, \dots, s_n\}$ represents the sensors, which are randomly deployed in W . All the sensors deployed are static and rechargeable. There is m number of mobile chargers employed to recharge the sensors. The value of $m (> 1)$ will be detailed in phase I of section IV. There is one stationary BS to support energy for all the mobile chargers. There are two threshold values designed for each sensor called first and second thresholds, denoted by η_1 and η_2 . The detailed calculation of the first threshold η_1 is given in phase II of section IV. Meanwhile, the second threshold η_2 is the energy required for the basic operation of the sensor. Let $T = \{t_1, t_2, \dots, t_j, \dots, t_l\}$ denote the total time observed for measuring the network surveillance quality (NSQ).

IV. THE PROPOSED ALGORITHM

There are three phases in the proposed ODRS algorithm. They are *Area Partitioning (AP) Phase*, *Threshold Value Calculation (TVC) Phase* and *Scheduling and Cooperation (SC) Phase*. The *AP phase* firstly partitions the network into m subregions based on the RL of the sensors and finds the best partitioning. Then the second phase determines the threshold value for each sensor to forward the request to its corresponding mobile charger. In the *SC phase*, the recharge scheduling of mobile chargers and cooperation between them is discussed.

A. Area Partitioning (AP) Phase

In the literature, most of the works statically partitioned the network. That is to say, the network was partitioned into equal-sized subregions and one mobile charger was assigned to each subregion. This type of partitioning is not very practical from the aspect of sensors RL. If the network is partitioned into equal-sized subregions regardless of the RL, the charging load of each mobile charger is different. Thus, few sensors may exhaust their energy because of not recharging timely.

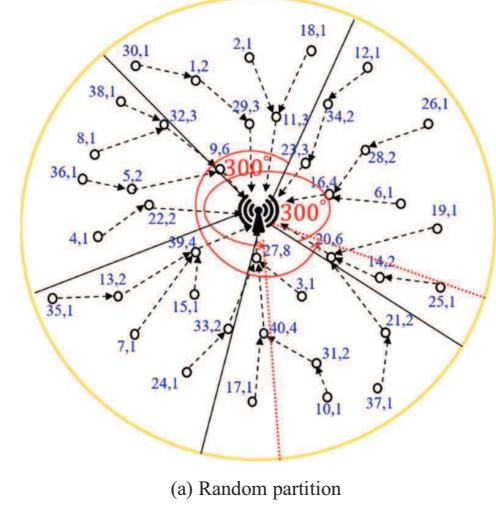
In large-scale sensors networks, the density (D) of the sensors is very high. If the density is high, the RL of the sensors is also high. If the RL is high, the energy consumption (EC) of the sensors is also high. That is, $D \propto RL \propto EC$. If the EC of the sensors is high, many sensors will forward the request to its mobile charger, which causes a higher charging load. Therefore, static partition leads to an imbalanced charging load for mobile chargers. Therefore, this paper

partitions the network based on the RL of the sensors. During the network initialization, the BS randomly partitions the area into m subregions, which are shown as black solid lines in Fig. 1(a). After partitioning into m subregions the RL of each subregion is calculated. Let N_L denote the total network load. Let RL_i denote the RL of the i^{th} subregion. That is,

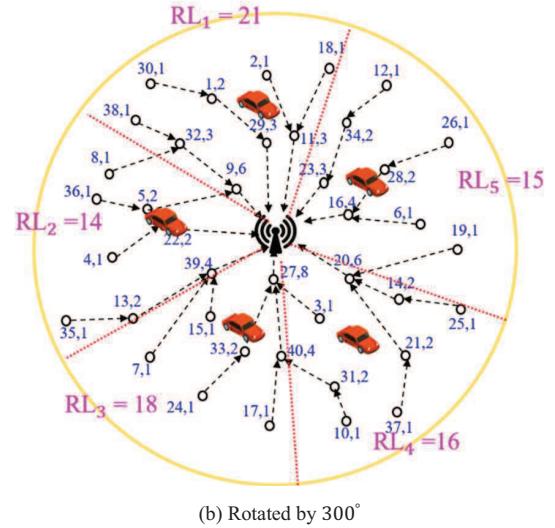
$$RL_i = \frac{N_L}{m} \quad (1)$$

Then all the partitions are rotated by $60^\circ, 120^\circ, 240^\circ, 300^\circ$ individually as four different cases. For the ease of presentation, only the fourth case where the random partitions rotated by 300° is shown in Fig. 1(a). The red dotted lines represent the 300° partition subregions. The final output of the fourth case is shown in Fig. 1(b). The RL is calculated for each partition in each case as shown in Fig. 1(b). To know the best partition, the load deviation of four cases should be calculated. Let RL_{avg}^{rnd} represents the average regional load of all regions in a random partitioning case. Let δ^{rnd} denote the load deviation of the random partition. The determination of δ^{rnd} is shown in (2).

$$\delta^{rnd} = \sum_{i=1}^m |RL_{avg}^{rnd} - RL_i| \quad (2)$$



(a) Random partition



(b) Rotated by 300°

Fig. 1. The considered network environment.

Similarly, δ^{60} , δ^{120} , δ^{240} and δ^{300} is also calculated. A partition with minimum load deviation is the best partition. After finding the best partition, the BS assigns one mobile charger for each subregion, where it is responsible for recharging the sensors in its corresponding region. The output of phase I is shown in Fig. 1(b), where the value of $m = 5$.

B. Threshold Value Calculation (TVC) Phase

As mentioned in section III, there are two threshold values for each sensor, i.e., η_1 and η_2 . Recall that η_2 is fixed for all the sensors. In the TVC phase, the detailed calculation of the η_1 is discussed. All the sensor requests are stored in the queue of the mobile charger, which is similar to the concept of queuing theory. Therefore, in this paper $M/G/1:/\infty/N$ queuing theory model is adopted to calculate the first threshold value η_1 . Let λ and μ denote the arrival and service rate of the requested sensors in the network. Therefore $M/G/1:/\infty/N$ can be expressed as $\lambda/\mu/1:/\infty/N$. Let L_q and L_s denote the average length of queue and system. Let W_q and W_s denote the average waiting time in queue and system. Assume, each mobile charger knows all the parameters L_q, L_s, W_q and W_s according to the formula calculation of queuing theory.

Let MC^k denote the k^{th} mobile charger. The operations of this phase are designed only for MC^k , which is the same for all the other mobile chargers in the network. Each MC^k use quorum to broadcast the schedule to its corresponding sensors. Let $L_q^{t,k}$ denote the queue size of k^{th} mobile charger at current time t . Let s_i^k denote the sensor s_i which belongs to MC^k . Assume sensor s_i^k wants to send the request to MC^k . Let γ variable denoting the total distance and v denote the speed of the mobile charger. Let T^{max} denote the time taken by the mobile charger to travel in its region. The notation T^{max} is calculated in (3).

$$T^{max} = \frac{\gamma}{v} \quad (3)$$

Let $e_{i,t}^{disch}$ denote the discharging rate of the sensor s_i at time t . Let $\eta_{i,1}^{k,t}$ denote the first threshold value of s_i^k at time t , which is shown in (4).

$$\eta_{i,1}^{k,t} = \eta_2 + (T^{max} * L_q^{t,k} + W_q^{t,k}) * e_{i,t}^{disc} \quad (4)$$

where $W_q^{t,k}$ denote the waiting time of k^{th} mobile charger at current time t . Thus, according to (4), all the sensors can calculate the η_1 value.

C. Scheduling and Cooperation (SC) Phase

In the SC phase, the MC^k constructs its path based on the BCQ of sensors. As shown in Fig. 2, during the network initialization at time t since there is no request in the queue, therefore $L_q^{t,k} = \{\emptyset\}$. Let $\wp_{current}^k$ and $\wp_{updated}^k$ denote the current path and updated path of MC^k . According to the calculation of (4), assume at the time t_1 , the MC^k receives the recharging requests from $\hat{s}_{23}, \hat{s}_{12}, \hat{s}_{19}, \hat{s}_6$. Initially, the MC^k constructed shortest Hamiltonian path for $\hat{s}_{23}, \hat{s}_{12}, \hat{s}_{19}, \hat{s}_6$. Whenever the MC^k moves to recharge the new sensor, $L_q^{t,k}$ and $W_q^{t,k}$ of corresponding time will be broadcasted using quorum. At time t_1 , when MC^k moves to sensor \hat{s}_{23} to recharge, the MC^k will broadcast $L_q^{t_1,k}, W_q^{t_1,k}$.

These two parameters $L_q^{t_1,k}, W_q^{t_1,k}$ are used to calculate $\eta_{i,1}^{k,t}$ for the sensors who did not send the request. At t_2 , MC^k is recharging \hat{s}_{12} and receives the recharging request from the \hat{s}_{26} . Assume BCQ of \hat{s}_{26} is higher than BCQ of \hat{s}_6 . Therefore, \hat{s}_{26} will be included in the path while \hat{s}_6 will be dropped out from $\wp_{current}^k$, aiming to achieve the highest NSQ. All the above-discussed procedure is shown in Fig. 2. Let \hat{s}_j^k denote the dropped out sensor. How to recharge the \hat{s}_j^k will be discussed in the following.

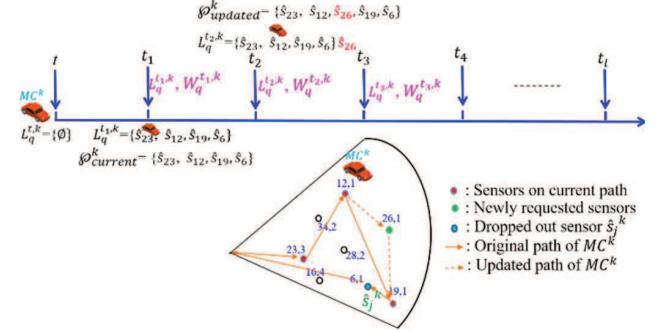


Fig. 2. Timeline schedule and example path of MC^k .

Since the charging load (the number of requests received) of each mobile charger is different, they can cooperate with the neighboring mobile chargers. According to the example shown in Fig. 2, the \hat{s}_j^k is \hat{s}_6 . The MC^k can communicate with the neighboring mobile chargers (MC^{k-1} and MC^{k+1}) to include \hat{s}_j^k to their schedule. Then both MC^{k-1}/MC^{k+1} includes \hat{s}_j^k to their schedule and checks whether the waiting time satisfies, according to (5) and \hat{s}_j^k is alive until MC^{k-1}/MC^{k+1} comes to recharges it, by satisfying (6). Let E_j^{rem} denote the remaining energy of \hat{s}_j^k . Let $T_{j,k-1}^{wait}$ denote the waiting time of \hat{s}_j^k for MC^{k-1} .

$$T_{j,k-1}^{wait} = \begin{cases} \tau_j^{mv}, & j = 1 \\ T_{j-1}^{wait} + \tau_j^{crg} + \tau_j^{mv}, & 2 \leq j \leq n \end{cases} \quad (5)$$

$$E_j^{rem} - e_j^{disch} * T_{j,k-1}^{wait} > 0 \quad (6)$$

According to (5) and (6), the calculation of $T_{j,k+1}^{wait}$ is similar for MC^{k+1} . Both the MC^{k-1} and MC^{k+1} calculate the $T_{j,k-1}^{wait}$ and $T_{j,k+1}^{wait}$. If both MC^{k-1} and MC^{k+1} satisfies (6), then the mobile charger which has a shorter waiting time for recharging \hat{s}_j^k will recharge \hat{s}_j^k . Let M^{best} denote the best mobile charger which takes care of \hat{s}_j^k .

$$M^{best} = \begin{cases} MC^{k-1}, & \text{if } T_{j,k+1}^{wait} > T_{j,k-1}^{wait} \\ MC^{k+1}, & \text{otherwise} \end{cases} \quad (7)$$

In the special case, if both MC^{k-1} and MC^{k+1} are too busy and does not satisfy (6), then both the MC^{k-1} and MC^{k+1} will abandon the request and notifies the MC^k . Assume $M^{best} = MC^{k-1}$ has a shorter waiting time $T_{j,k-1}^{wait}$. Therefore, the MC^{k-1} includes \hat{s}_j^k to its schedule and recharges \hat{s}_j^k .

V. EXPERIMENTAL EVALUATIONS

The proposed ODRS algorithm is compared with the existing NJNP [6], ALG [8], GA [9] DWDP [11] and MCCA [12]. All the compared studies employed mobile chargers to recharge the sensors. The NJNP recharges the sensors based on the nearest distance while ALG considered the parameters such as network density, distance and remaining energy. The GA algorithm adopted the greedy and genetic algorithms to schedule the mobile charger. The DWDP algorithm recharges the sensors by considering the spatial and temporal priority. Besides, the MCCA algorithm constructed a multi-charger cooperation algorithm for uneven clusters.

A. Environment

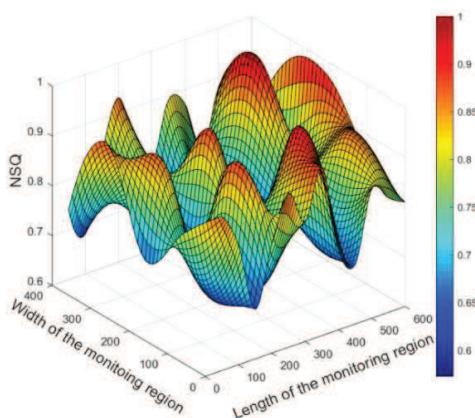
The simulation tool used is MATLAB R2019b. The area size is $600m \times 400m$. The deployed sensors are 500 while the sensing range is set at $10m$. The value of η_2 is set from 0.1J to 0.7J. The value of e_i^{disch} is 0.05J/s. The recharging rate of the mobile charger is 10 J/s. The value of v is set at 5 m/s. The parameters are given in Table I.

TABLE I. PARAMETERS

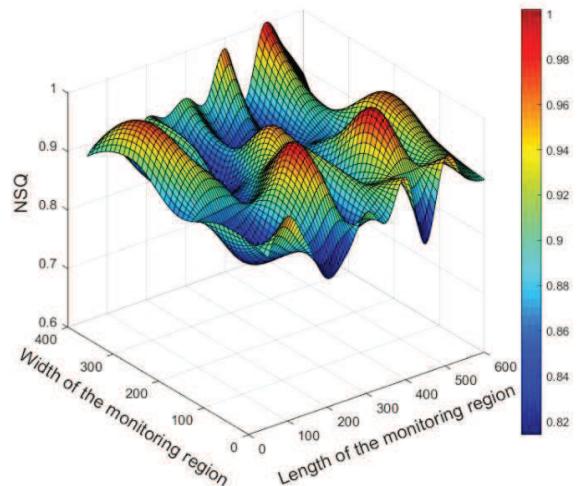
Parameters	Values
Tool	MATLAB R2019b
Area size	$600m \times 400m$
Number of sensors	500
Deployment type	Random
Sensing range	$10m$
η_2	0.1 – 0.7J
e_i^{disch}	0.05J/s
v	5 m/s
Recharging rate	10 J/s

B. Results

Figs. 3(a) and 3(b) compares the NSQ of the network with and without considering the BCQ of the requested sensors. In this experiment, the length and width of the monitoring region is $600m \times 400m$. The sensing range of sensors is set at $10m$. In Fig. 3(a), the lowest and highest NSQ of the network is 0.62 and 0.972, respectively. This result is obtained without considering the BCQ of the recharging requested sensors. All the sensors are recharged according to the sequence in the queue.



(a). Without considering the BCQ of the requested sensors.

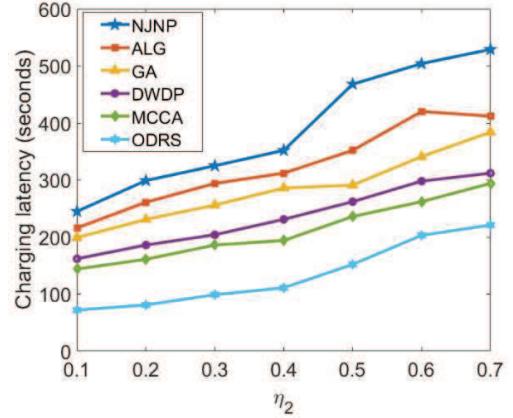


(b). With considering the BCQ of the requested sensors (proposed ODRS)

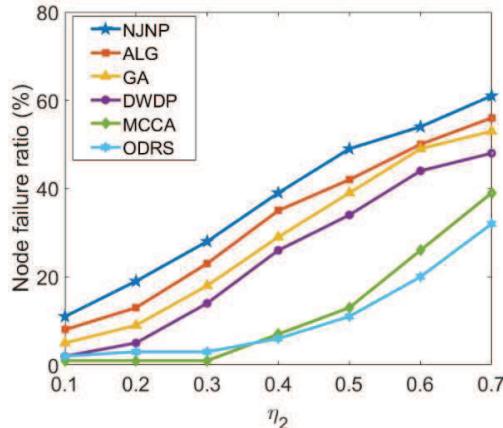
Fig. 3. The NSQ with and without considering BCQ of the requested sensors.

On the other hand, in Fig. 3(b), the mobile charger recharges the requested sensors based on the BCQ. As a result, the lowest and highest NSQ obtained in Fig. 3(b) is 0.823 and 1, respectively. In comparison, the NSQ of Fig. 3(a) is not stable and it is lower compared to Fig. 3(b).

Figs. 4(a) and 4(b) compares the charging latency and node failure ratio by varying the value of η_2 . The value of η_2 is varied from 0.1 to 0.7. As shown in Fig. 4(a), the charging latency increases with η_2 . This occurs because if the value of η_2 increases, the sensor will go to a sleep state very quickly. Therefore, many sensors will send the recharging requests. This makes the mobile charger to be very busy leading to higher charging latency. In comparison, the proposed ODRS algorithm yields the lowest charging latency. This occurs because the proposed ODRS algorithm considered the cooperation between neighboring mobile chargers. Therefore, the sensors are recharged very quickly, leading to the lowest charging latency. The charging latency of MCCA is also lower since it employed multiple mobile chargers. The other four algorithms considered parameters such as distance, remaining energy, density, etc. Therefore few sensor nodes requests are ignored, leading to the higher charging latency.



(a). The value of η_2 Vs charging latency



(b). The value of η_2 Vs node failure ratio

Fig. 4. Comparison of charging latency and node failure ratio by varying η_2 .

In Fig. 4(b), the node failure ratio is compared for the four algorithms by varying the value of η_2 . The node failure ratio is increased with the value of η_2 . This occurs because if the value of η_2 increases, the charging load is very high. Therefore, the requested sensors whose BCQ is low will be ignored. Thus, the node failure ratio is increased. In comparison, the proposed ODRS algorithm yields the lowest node failure ratio. This is due to the cooperation between mobile chargers. The NJNP yields the highest node failure ratio because it ignored the requests of the sensors which are far away from the mobile charger.

VI. CONCLUSIONS

This paper proposed an on-demand recharge scheduling (ODRS) algorithm. The proposed ODRS algorithm constructs the path by considering the BCQ of the requested sensors, aiming to maximize the NSQ. There are three phases in the proposed ODRS algorithm: including *Area Partitioning (AP) Phase*, *Threshold Value Calculation (TVC) Phase* and *Scheduling and Cooperation (SC) Phase*. The AP phase partitions the network into m subregions according to the RL of the sensors. The TVC phase discussed the threshold value of each sensor based on the real-time information of the mobile charger. The final phase presents the schedule of the

mobile charger and also discussed the cooperation between neighboring mobile chargers. The future work of this study is to predict the energy consumption rate of sensors and energy transfer between mobile chargers.

REFERENCES

- [1] T. Liu, B. Wu, S. Zhang, J. Peng and W. Xu, "An effective multi-node charging scheme for wireless rechargeable sensor networks," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications, 2020, pp. 2026-2035.
- [2] G. Jiang, S.-K. Lam, Y. Sun, L. Tu and J. Wu, "Joint charging tour planning and depot positioning for wireless sensor networks using mobile chargers," IEEE/ACM Transactions on Networking, vol. 25, no. 4, pp. 2250-2266, Aug. 2017.
- [3] A. Tomar, K. Nitesh and P.K. Jana, "An efficient scheme for trajectory design of mobile chargers in wireless sensor networks," Wireless Networks, vol. 26, pp. 897-912, Feb. 2020.
- [4] N. Gharaei, Y. D. Al-Otaibi, S. A. Butt, S. J. Malebary, S. Rahim and G. Sahar, "Energy-efficient tour optimization of wireless mobile chargers for rechargeable sensor networks," IEEE Systems Journal, vol. 15, no. 1, pp. 27-36, March 2021.
- [5] R. Anwit, A. Tomar and P.K. Jana, "Tour planning for multiple mobile sinks in wireless sensor networks: A shark smell optimization approach," Applied Soft Computing, vol. 97, pp. 106802, Dec. 2020.
- [6] L. He, L. Kong, Y. Gu, J. Pan and T. Zhu, "Evaluating the on-demand mobile charging in wireless sensor networks," IEEE Transactions on Mobile Computing, vol. 14, no. 9, pp. 1861-1875, 1 Sept. 2015.
- [7] B. Dande, S. -Y. Chen, H. -C. Keh, S. -J. Yang and D. S. Roy, "Coverage-aware recharging scheduling using mobile charger in wireless sensor networks," IEEE Access, vol. 9, pp. 87318-87331, June. 2021.
- [8] A. Kaswan, A. Tomar and P.K. Jana, "An efficient scheduling scheme for mobile charger in on-demand wireless rechargeable sensor networks," Journal of Network and Computer Applications, vol. 114, pp. 123-134, July. 2018.
- [9] Z. Wei, M. Li, Q. Zhao, Z. Lyu, S. Zhu and Z. Wei, "Multi-MC charging schedule algorithm with time windows in wireless rechargeable sensor networks," IEEE Access, vol. 7, pp. 156217-156227, Oct. 2019.
- [10] H. Yu, C. -Y. Chang, Y. Wang, D. S. Roy and X. Bai, "CAERM: Coverage aware energy replenishment mechanism using mobile charger in wireless sensor networks," IEEE Sensors Journal, vol. 21, no. 20, pp. 23682-23697, Oct. 2021.
- [11] C. Lin et.al., "Double warning thresholds for preemptive charging scheduling in wireless rechargeable sensor networks," Computer Networks, vol. 148, pp. 72-87, Jan. 2019.
- [12] G. Han, H. Wang, H. Guan and M. Guizani, "A mobile charging algorithm based on multi charger cooperation in Internet of Things," IEEE Internet of Things Journal, vol. 8, no. 2, pp. 684-694, Jan. 2021.

Mobile Charger Scheduling Algorithm for Energy Recharging in Wireless Sensor Networks

Bhargavi Dande*

Corresponding author

Dept. of Computer Science and
Information Engineering
Tamkang University
New Taipei City, Taiwan
808415011@gms.tku.edu.tw

Chih-Yung Chang

Dept. of Computer Science and
Information Engineering
Tamkang University
New Taipei City, Taiwan
cychang@mail.tku.edu.tw

Cheng-De Fan

Dept. of Electrical and
Computer Engineering
Tamkang University
New Taipei City, Taiwan
408500477@gms.tku.edu.tw

Abstract— In wireless sensor networks (WSNs), the finite battery capacity of the sensor nodes restricts the lifetime of the network. With the wireless power transfer (WPT) technology, the mobile chargers can transmit energy to the sensor nodes, which provides a solution for recharging the sensors. Therefore, this paper proposed a mobile charger scheduling (MCS) algorithm, which maximizes the coverage quality of the network. The proposed MCS algorithm formulated an adaptive recharging request threshold for the sensor nodes. Then the coverage benefit (CB) of the requested sensors is considered to dynamically adjust the path. Finally, through the experimental analysis, it is shown that the proposed MCS algorithm gains the best performance than the existing works.

Keywords— *sensors, mobile charger, wireless rechargeable sensor networks, wireless power transfer.*

I. INTRODUCTION

Wireless sensor networks (WSNs) are comprised of tiny sensors, which have been widely used in many real applications. Even though the sensors were used in many applications, the limited battery capacity is a big issue in WSNs. Therefore, energy recharging the sensors is very important in WSNs. In the recent literature, many algorithms used wireless power transfer (WPT) technology to recharge the sensors.

The existing studies, which adopt the WPT technology for recharging, were divided into two types: offline mechanisms [1] – [5] and online mechanisms [6] – [12]. In the first type, the mobile charger travels in a predefined trajectory and the sensors were recharged periodically. These types of mechanisms ignored that the energy discharging rate is different for the sensors in the network. On the other hand, the second type of mechanism allows the sensors to send the real-time recharging request. In the online mechanisms, there is a threshold value for each sensor battery. When the sensor energy meets the threshold value, the sensor will initiate a request. Thus, the mobile charger designs its recharging path to recharge the sensors. There might be many recharging requests received to the mobile charger, choosing the best recharging candidate is a very important task in online mechanisms. An online MCS algorithm maximizes the coverage quality of the network. An adaptive threshold value and insertion and deletion of sensors to the path are designed in this algorithm.

II. RELATED WORKS

The literature on the offline and online mechanisms is presented in this section.

A. Offline Mechanisms

Study [1] aimed to reduce the mobile chargers used and investigated the cooperation between them. Study [2] constructed the path, to reduce the sensor dead duration. Ma et al. [3] assumed all the sensors that fall in the recharging range were recharged. This study aims to reduce the path length by recharging multiple sensors.

Lyu et al. [4] adopted a genetic algorithm for constructing the path of the mobile charger. This study considered that the energy discharging rate of sensors is the same, which is not suitable for many real applications. Hwang et al. [5] assumed that the sensors were partially recharged in each charging round. This study adopted the genetic algorithm, to shorten the dead nodes. However, partial recharging mechanisms consume a higher amount of energy for the movement. The energy discharging rate of sensors might differ in WSNs, which is very difficult to deal with offline mechanisms.

B. Online Mechanisms

Lin et al. [6] adopted the game theory to investigate the charging issues. Wang et al. [7] aimed to reduce the sensor node dead time duration. This study assumed that the sensors were partially recharged. Besides, the sensors were prioritized based on their monitoring tasks contribution. However, the partial charging schemes lead to the ping-pong movement of the mobile charger.

Study [8], a double warning threshold value for each sensor is designed. Tomar et al. [9], adopted the fuzzy logic to investigate the on-demand charging for sensor nodes. This study predefined the threshold value of each sensor. The parameters considered for constructing the recharging path are the remaining energy of sensor nodes and node density. However, the coverage benefit (CB) of the requested sensors is ignored.

Kaswan et al. [10] scheduled the mobile charger based on the linear programming formulation and gravitational search algorithm. Feng et al. [11] adjusted the charging duration of sensors to reduce the node failure ratio. Zhao et al. [12] investigated the service cost to schedule the mobile charger. This study proposed a dynamic insertion algorithm to improve the charging efficiency. Although the studies [6] – [12], dynamically constructed the recharging path of the mobile charger, none of them considered the CB of the requested sensors. That is to say, the coverage quality of the network is ignored completely.

In the proposed mobile charger scheduling (MCS) algorithm, an adaptive recharging request is designed for each sensor. To issue the recharging request, each sensor

determines an appropriate threshold value by considering its discharging rate and total time taken by the mobile charger.

III. NETWORK ENVIRONMENT

A WSNs, which consists of n static homogeneous sensors, represented by $S = \{s_1, s_2, \dots, s_n\}$ is considered in this paper. Let \mathcal{R} denote the considered monitoring region. All the sensors are randomly deployed in \mathcal{R} . The mobile charger recharges the sensor nodes. There is one stationary sink located at the center of \mathcal{R} , which is responsible for supporting the energy to the mobile charger. Let E_S^{max} represents the maximum battery of sensor nodes. Let e_i^{dis} denote the discharging rate of sensor nodes. This paper assumes that each sensor has two threshold values called *request threshold* and *minimum energy threshold*, which are denoted by $e_{th,i}^{req}$ and e^{min} , respectively. The value of e^{min} is predefined for all the sensors while the detailed calculation of $e_{th,i}^{req}$ is discussed in section IV. Let $T = \{t_1, t_2, \dots, t_j, \dots, t_m\}$ denote the total time for measuring the coverage quality of the network.

IV. THE PROPOSED ALGORITHM

The MCS algorithm comprises two phases, namely *Request Threshold Calculation (RTC) Phase*, and *Path Construction (PC) Phase*. The *RTC* phase aims to calculate the threshold value of each sensor for sending the recharge request. In the second phase, the path is constructed by considering additional requests from the service pool ξ and dynamically reconstructs a new recharging path.

A. Request Threshold Calculation (RTC) Phase

In most of the existing studies, the threshold value of each sensor is fixed. That is to say, if the remaining energy of any sensor reaches a certain value, the sensor will give a request. This policy leads to many issues in the network. A threshold value should not be too tight or too loose. A tight threshold value means the sensor initiates the request when it has lower remaining energy. This leads to sensor energy exhaustion by the time the mobile charger reaches the sensor. On the other hand, a loose threshold value means, the sensor initiates the request when it has more remaining energy. Therefore, *request threshold calculation* is very important in online mechanisms.

Let \tilde{S} represent the sensors, which need to send a recharging request. Assume sensor $\tilde{s}_i \in \tilde{S}$ aims to calculate its threshold value for sending the recharge request. The proposed MCS algorithm considered the discharging rate e_i^{disch} of sensor $\tilde{s}_i \in \tilde{S}$ and the length of the service pool ξ to calculate the *request threshold* $e_{th,i}^{req}$ of each sensor $\tilde{s}_i \in \tilde{S}$. In this paper, the mobile charger uses quorum to broadcast its current schedule to all the sensors periodically. Assume there are u requests stored in ξ represented as $\{\hat{s}_1, \hat{s}_2, \dots, \hat{s}_u\}$. Let l_u denote the location of the sensor \hat{s}_u . Assume that the current recharge schedule of the mobile charger is represented as $((\hat{s}_1, l_1), (\hat{s}_2, l_2), \dots, (\hat{s}_v, l_v))$.

To calculate the threshold value $e_{th,i}^{req}$ of each sensor \tilde{s}_i , the total time taken by the mobile charger should be calculated. The total time includes the time taken for moving and recharging each sensor in ξ . Let e_M^{charg} denote the recharging rate of the mobile charger. Let t_u^{crg} means the time taken to fully recharge \hat{s}_u , which is shown in (1).

$$t_u^{crg} = \frac{E_S^{max}}{e_M^{charge}} \quad (1)$$

Let v denote the mobile charger moving speed. Let t_u^{mv} represents the time taken by the mobile charger to move from sink/ \hat{s}_{k-1} to \hat{s}_k . Let \hat{s}_0 denote the sink node.

$$t_u^{mv} = \frac{d(\hat{s}_{k-1}, \hat{s}_k)}{v}, 1 \leq k \leq u \quad (2)$$

By combining the charging and moving times of each sensor in ξ , the total time required for the sensor $\tilde{s}_i \in \tilde{S}$ to wait for the mobile charger can be calculated. Let τ_M^{need} represent the total time taken by the mobile charger to recharge and move to all the sensors in $\xi : \{\hat{s}_1, \hat{s}_2, \dots, \hat{s}_u\}$. Assume sensor $\tilde{s}_i \in \tilde{S}$ is the next requesting sensor whose request is stored in the $(u+1)^{th}$ position. The value of τ_M^{total} can be calculated as shown in (3).

$$\tau_M^{total} = \sum_{k=1}^u t_u^{crg} + \sum_{k=1}^u t_u^{mv} \quad (3)$$

Finally, the threshold value of the sensor $\tilde{s}_i \in \tilde{S}$ is evaluated as shown in (4).

$$e_{th,i}^{req} = e^{min} + \tau_M^{total} * \delta * e_i^{disch} \quad (4)$$

where δ is constant and e_i^{disch} is the discharging rate of the sensor \tilde{s}_i . Therefore, each sensor $\tilde{s}_i \in \tilde{S}$ can determine its threshold value $e_{th,i}^{req}$. The request packet format can be represented as $<id, l_i, E_{i,j}^{rem}, e_i^{disch}, t_j>$, where id denotes the sensor id, l_i is location, $E_{i,j}^{rem}$ is the remaining energy, and t_j is the current time slot, respectively.

B. Path Construction (PC) Phase

In the *PC* phase, the mobile charger dynamically reconstructs its path by considering the coverage benefit (CB) of newly requested sensors in ξ . Assume, at the current time t , sensor $\tilde{s}_m \in \tilde{S}$, has sent the recharging request. If the requested sensor $\tilde{s}_m \in \tilde{S}$ in ξ has the largest CB compared to the CB of any sensor on the current path $\wp^{current}$, the $\tilde{s}_m \in \tilde{S}$ will be inserted to $\wp^{current}$ while the sensor with the least CB on $\wp^{current}$ will be deleted from $\wp^{current}$ and send to ξ . The CB calculation of $\tilde{s}_m \in \tilde{S}$ is presented in the following.

Let N_m represents the set of neighbor nodes of the sensor \tilde{s}_m . Let β_m and β_{N_m} represents the coverage area of the sensor \tilde{s}_m and its neighbors, respectively.

$$\beta_{N_m} = \bigcup_{s_j \in N_m} \beta_j \quad (5)$$

Let $\zeta_{j,t}^{wrk}$ denote a Boolean variable denoting sensor s_j is working/sleeping at the current time t .

$$\zeta_{j,t}^{wrk} = \begin{cases} 1, & E_{j,t}^{rem} > e^{min} \text{ (working)} \\ 0, & E_{j,t}^{rem} < e^{min} \text{ (sleeping)} \end{cases} \quad (6)$$

Let $c_{N_m}^t$ denote the CB of N_m at current time t , which can be derived according to (7).

$$c_{N_m}^t = \bigcup_{s_j \in N_m} (\beta_j * \zeta_{j,t}^{wrk}) \quad (7)$$

Let c_m^t denote the CB of \tilde{s}_m at current time t . That is,

$$c_m^t = (\beta_m - c_{N_m}^t) * \zeta_{m,t}^{wrk} \quad (8)$$

Let B_m denote the CB of \tilde{s}_m . If \tilde{s}_m is included in the path, the path length might be increased. Therefore, path

length should be calculated first. Let Δs_m denote the increased path length. The notation Δs_m is calculated as shown in the (9).

$$\Delta s_m = d_{i,m} + d_{m,i+1} - d_{i,i+1} \quad (9)$$

$$B_m = \frac{c_m^t}{\Delta s_m} \quad (10)$$

To present the operations of this phase, a *passer-by update* (*PU*) strategy will be introduced. In the insertion of *PU* strategy, at most one sensor from the service pool ξ to $\wp^{current}$ is included. Similarly, in the deletion of *PU* strategy, at most one sensor is deleted from $\wp^{current}$ and send to ξ .

Let B_m^ξ and B_i^\wp denote the CB of each sensor $\hat{s}_m \in \xi$ and $s_i^\wp \in \wp^{current}$, respectively. Let S_{best}^ξ denote the sensor with the largest CB is B_m^ξ in ξ . That is,

$$S_{best}^\xi = \arg \max_{s_m \in \xi} B_m^\xi \quad (11)$$

Let S_{worst}^\wp denote the sensor with the least CB is B_i^\wp in $\wp^{current}$.

$$S_{worst}^\wp = \arg \min_{s_i^\wp \in \wp^{current}} B_i^\wp \quad (12)$$

Let B_{avg}^ξ and B_{avg}^\wp denote the average CB of the sensors in ξ and $\wp^{current}$, respectively. The value of B_{avg}^ξ and B_{avg}^\wp is expressed in (13) and (14).

$$B_{avg}^\xi = \left(\sum_{m=1}^{|\xi|} B_m^\xi \right) / |\xi| \quad (13)$$

$$B_{avg}^\wp = \left(\sum_{i=1}^{|\wp|} B_i^\wp \right) / |\wp| \quad (14)$$

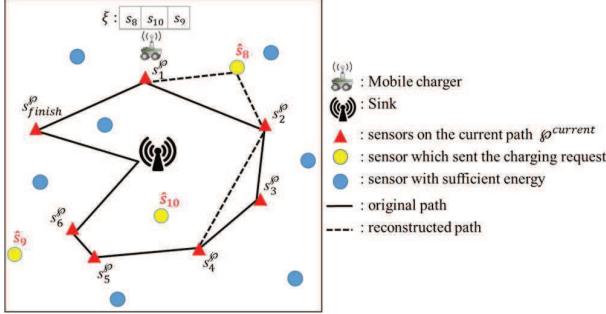


Fig. 1. The path is constructed by adopting the *PU* strategy.

Let B_{best}^ξ and B_{worst}^\wp denote the CB of S_{best}^ξ and S_{worst}^\wp , respectively. The insertion and deletion conditions of each sensor $\hat{s}_m \in \xi$ and $s_i^\wp \in \wp^{current}$ of *PU* strategy is presented in the following. Fig. 1 depicts an example of the *PU* strategy.

Insertion of PU strategy:

$$B_{best}^\xi > B_{avg}^\wp \quad (15)$$

Deletion of PU strategy:

$$B_{worst}^\wp < B_{avg}^\xi \quad (16)$$

As shown in Fig. 1, $\wp^{current} = \{s_{fini}^\wp, s_1^\wp, s_2^\wp, s_3^\wp, s_4^\wp, s_5^\wp, s_6^\wp\}$. Assume that the mobile charger completed recharging s_{fini}^\wp and recharging the sensor s_1^\wp at current

time t . During the recharging mission at s_1^\wp , the mobile charger receives the request from \hat{s}_8 , \hat{s}_{10} and \hat{s}_9 . To include the best sensor S_{best}^ξ to $\wp^{current}$, the mobile charger calculate the CB of the sensors \hat{s}_8 , \hat{s}_{10} and \hat{s}_9 according to (8). Since the CB of \hat{s}_8 is higher than s_3^\wp , the sensor $S_{best}^\xi = \hat{s}_8$ is inserted into the path while $S_{worst}^\wp = s_3^\wp$ is deleted from $\wp^{current}$.

V. EXPERIMENTAL EVALUATIONS

This section compares the simulation results of the proposed MCS with the ETLBO, Mersh and Earliest Deadline First (EDF) algorithms. The Mersh algorithm [11] constructs the path by considering the tolerable latency of the requested sensors. The ETLBO [12] proposed charging scheduling by formulating the best candidate position and time to insert the new charging candidates to the original path. Finally, the EDF algorithm recharges the sensors based on the earliest deadline.

A. Simulation Environment

The simulation tool used is MATLAB R2019b. The area size is $600m \times 400m$. The deployed sensors are 400 to 700. The sensing range is set at $5m$ – $20m$. The e^{min} is set from $0.4J$ to $0.7J$. The value of e_i^{disch} is set at $0.05J/s$. E_S^{max} is set at $3.6kJ$. The e_M^{charg} is adjusted between $1J/s$ to $15 J/s$. The value of v is varied ranging from 1 to $6 m/s$. Table I presents all the simulation parameters.

TABLE I. SIMULATION PARAMETERS

Parameters	Values
Simulator	MATLAB R2019b
Monitoring area	$600m \times 400m$
Number of deployed sensors	400 - 700
Deployment type	Random
Sensing range of sensors	5 – $20m$
e^{min}	0.4 - 0.7J
e_i^{disch}	0.05J/s
E_S^{max}	3.6kJ
v	1 - $6 m/s$
e_M^{charge}	1-15 J/s

B. Simulation Results

Fig. 2 evaluates the coverage qualities of the MCS algorithm with the other three algorithms. Since the number of deployed sensors increases the density of the network increases, leading to higher quality. On the other hand, the coverage quality increases with the sensing range. This is due to a larger sensing range covering a large area. Among the four algorithms, the proposed MCS outperforms the coverage quality. Besides, both the *ETLBO* and *Mersh* algorithms improved the charging efficiency and reduced the dead nodes while the *EDF* algorithm was always considered the most urgent sensor. However, none of the three algorithms considered the CB of recharging requested sensors, leading to the lower coverage qualities.

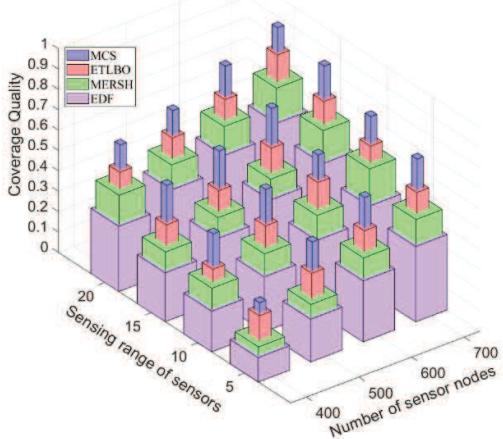


Fig. 2. The coverage quality comparison for four algorithms.

Fig. 3 compares the recharged sensor's contribution by varying speed of mobile charger v . The value of v is set from 1 to 6 (m/s). As shown in Fig. 3, the recharged sensors contribution of the proposed MCS increases with v . This is because the mobile charger can recharge many sensors when it travels at a higher speed. Thus, it increases the recharged sensor's contribution. On the other hand, for the other three algorithms, the recharged sensor contributions increase and decrease regardless of the value of v . This is due to ignoring the CB of the requested sensors.

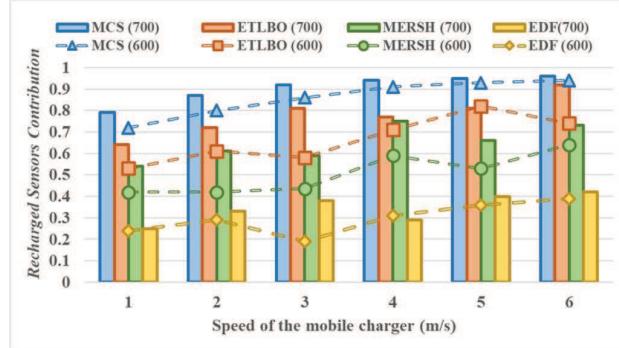


Fig. 3. Comparison of recharged sensors contribution.

Fig. 4 further compares the coverage loss by varying the e_M^{charge} and e^{min} . The value of e_M^{charge} and e^{min} are varied ranging from 1 to 15 (J/s) and 0.4 to 0.7, respectively. As shown in Fig. 4, the coverage loss is decreased with the increasing recharging rate. This is because a higher recharging rate can speed up the recharging time. Therefore, most of the requested sensors can be recharged, reducing the coverage loss. On the other hand, the coverage loss is decreased with the minimum energy threshold value e^{min} . This is due to higher e^{min} value leads the sensors to fall into the sleep state very quickly. As shown in Fig. 4, the proposed MCS achieves a lower coverage loss compared to the other three algorithms. The proposed algorithm considered the CB of the requested sensors. Besides, the other three algorithms ignored the CB.

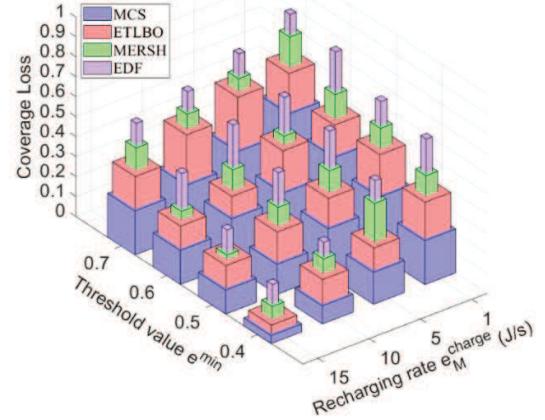


Fig. 4. Comparison of coverage loss by varying recharging rate and the threshold value.

VI. CONCLUSIONS

This paper proposed a mobile charger scheduling algorithm, namely MCS for the WSNs. The proposed MCS dynamically reconstructs its path based on the CB of the recharging requested sensors. The proposed algorithm consists of two phases: *RTC* and *PC Phases*. The *RTC* phase formulated an adaptive threshold value for the sensor nodes. In the second phase, the CB is of each recharging requested sensor is considered to construct its recharging path. In addition, a *passer-by update (PU)* strategy is proposed to insert the additional requests to the current path based on their CB. Our future work is to consider multiple mobile chargers and cooperation between them.

REFERENCES

- [1] Z. Chen, X. Chen, D. Zhang, and F. Zeng, "Collaborative mobile charging policy for perpetual operation in large-scale wireless rechargeable sensor networks," Neuro computing, vol. 270, pp.137–144, Dec. 2017.
- [2] W. Xu, W. Liang, X. Jia, Z. Xu, Z. Li and Y. Liu, "Maximizing sensor lifetime with the minimal service cost of a mobile charger in wireless sensor networks," IEEE Transactions on Mobile Computing, vol. 17, no. 11, pp. 2564-2577, Nov. 2018.
- [3] Y. Ma, W. Liang and W. Xu, "Charging utility maximization in wireless rechargeable sensor networks by charging multiple sensors simultaneously," IEEE/ACM Transactions on Networking, vol. 26, no. 4, pp. 1591-1604, Aug. 2018.
- [4] Z. Lyu, Z. Wei, J. Pan, H. Chen, C. Xia, J. Han, and L. Shi, "Periodic charging planning for a mobile WCE in wireless rechargeable sensor networks based on hybrid PSO and GA algorithm," Appl. Soft Comput., vol. 75, pp. 388 – 403, Feb. 2019.
- [5] T. T. Huong, P. L. Nguyen, H. T. T. Binh, K. Nguyenz, N. M. Hai and L. T. Vinh, "Genetic algorithm-based periodic charging scheme for energy depletion avoidance in WRSNs," 2020 IEEE Wireless Communications and Networking Conference (WCNC), 2020, pp. 1-6.
- [6] C. Lin *et al.*, "GTCCS: A game theoretical collaborative charging scheduling for on-demand charging architecture," IEEE Transactions on Vehicular Technology, vol. 67, no. 12, pp. 12124-12136, Dec. 2018.
- [7] K. Wang, L. Wang, M. S. Obaidat, C. Lin and M. Alam, "Extending network lifetime for wireless rechargeable sensor network systems through partial charge," IEEE Systems Journal, vol. 15, no. 1, pp. 1307-1317, March. 2021.
- [8] C. Lin, Y. Sun, K. Wang, Z. Chen, B. Xu, and G. Wu, "Double warning thresholds for preemptive charging scheduling in wireless rechargeable sensor networks," Computer Networks, vol. 148, pp. 72-87, Jan. 2019.

- [9] A. Tomar, L. Muduli, and P. K. Jana, "An efficient scheduling scheme for on-demand mobile charging in wireless rechargeable sensor networks," *Pervas. Mobile Comput.*, vol. 59, pp. 101074, Oct. 2019.
- [10] A. Kaswan, A. Tomar, and P. K. Jana, "An efficient scheduling scheme for mobile charger in on-demand wireless rechargeable sensor networks," *J. Netw. Comput. Appl.*, vol. 114, pp. 123 - 134, Jul. 2018.
- [11] Y. Feng, L. Guo, X. Fu and N. Liu, "Efficient mobile energy replenishment scheme based on hybrid mode for wireless rechargeable sensor networks," *IEEE Sensors Journal*, vol. 19, no. 21, pp. 10131-10143, Nov. 2019.
- [12] C. Zhao, H. Zhang, F. Chen, S. Chen, C. Wu, and T. Wang, "Spatiotemporal charging scheduling in wireless rechargeable sensor networks," *Comput. Commun.*, vol. 152, pp. 155 - 170, Feb. 2020.

UTILIZATION OF FOG COMPUTING IN TASK SCHEDULING AND OFFLOADING: MODERN GROWTH AND FUTURE CHALLENGES

1st Archana R.

Dept. of Computing Technologies

School of Computing,

SRMIST

Chennai, India.

ar6036@srmist.edu.in

2nd Pradeep Mohan Kumar,

Dept. of Computing Technologies

School of Computing

SRMIST

Chennai, India.

pradeepk@srmist.edu.in

Abstract—Recent advancements towards emerging technologies like the Internet of Things (IoT) needs resource-based processing for executing diverse real-time applications. In this IoT environment, the connected devices generate an enormous amount of data and processed in the cloud environment due to its service provisioning nature and device-scalability characteristics. Moreover, the request generated by the IoT devices towards the cloud is not effectively resolved; specifically, monitoring of resource availability, resource management, and task scheduling. The major challenge in the FC environment is task scheduling and resource allocation. The researches on FC is determined to be in its infancy nature where fog infra-structure, application, working environment, taxonomy-based allocations are still known the wider field in the research area. This work provides an extensive survey towards the synthesize of FC in research and industrial community to fulfil the need of users/researchers. Here, FC definitions, computation process, research trends and technical evaluations among the nodes are reviewed extensively. Similarly, the architectures, components are also discussed. The fog components are defined, and the nodes are deployed over the simulation environment. Similarly, the existing research gaps are also analyzed, and the gaps are bridged by examining the deficiency factors related to resource allocation and task scheduling. The limitations of FC with present research works are examined based on the open issues. This paves the way to proceed with future research enhancements over FC paradigm.

Keywords—Fog computing, resource allocation, task scheduling, resource management, computational process

I. INTRODUCTION

This The extensive adoption of smart devices and computers on various organizations helps to achieve daily tasks [1]. These adopted devices generate data through the sensors located and the applications used. As an outcome, the organization pretends to store and generate an enormous amount of data over a regular basis. With the IoT devices proliferation, the generated data with the sensors located are increased extensively. With the enormous increase in data, the volume is produced, and the traditional databases lack in processing the unstructured and structured data [2]. This leads to the wider growth of big data analytics. Every organization is now prioritizing the collected data analysis to haul out essential perspectives to make appropriate decisions [3]. In recent days, these industries require

dynamic IT infrastructure due to the shift in CC owing to its scalability, accessibility, and pay/use features. The common services offered by CC are with the use of Platform as a Service (PaaS), Software as a Service (SaaS), and Infrastructure as a Service (IaaS) [4]. Moreover, generated data from the sensors are known as big data which is not processed and transferred to the cloud.

In Fog computing (FC), the computing nature adopts decentralized process, which does not rely on centralized components such as Cloud Computing (CC) [5]. It has the competency to get rid of extensive latency of cloud with the usage of idle resources which is located nearer to the users. Moreover, FC relies on the cloud to carry out complex processing nature. FC is a decentralized computing nature, where CC does not. Various nodes are connected with FC paradigms which have the computational ability. At present, some devices with low power specifications possess processing capability of various cores [6]. Therefore, various devices such as network management devices, base stations, switches, routers, and smartphones equipped with storage capability and processing power which functions as Fog devices. The resources are idle for a certain period.

Moreover, the above-mentioned investigations have not discussed the taxonomy of infrastructural, applications, and platform requirements in FC. However, these authors do not show any interest in resource management, fault tolerance, task scheduling, and FC services. This work considers these above-mentioned issues and discusses these metrics in detail about CC. The significant contributions of this comprehensive review are given below:

- ✓ This work discusses the recent research trends in FC.
- ✓ Discusses various FC architectural model and higher-level architectures.
- ✓ Discusses the requirements essential for infrastructure deployment, applications, and platform for deployment.
- ✓ Construct the bridge to fulfil the research gap identified in previous research works. Metrics like task scheduling, resource management, and fault tolerance are discussed.

- ✓ Addresses the limitations encountered in recent research works and fulfil the open issues like platform, infrastructures and applications.

From this extensive review, the research community will have a deeper insight towards the preliminary requirements of modelling the FC paradigm with a better understanding of the fog needs. The remainder of this review is structured as section 2 is background studies, section 3 is fog definitions, section 4 is Fog architecture, section 5 is resource management, section 6 is task scheduling, section 7 is load balancing and offloading strategies, section 8 is a research platform. Section 9 is research directions, and section 10 is the conclusion.

A. Background

The researchers of Cisco systems define fog Computing (FC) term. Data processing and application logic is not a newer idea to be discussed along with FC. Cloudlets and FC are the advancements which evolve over the computing process. It is applied to connect various devices like IoT, which plays an essential role in Fog, CC, and IoT technologies, respectively. Similarly, fog is a non-virtualized and virtualized technology that offers storage, networking, and computational services like IoT devices and cloud servers. Moreover, these are not completely over the network edges [13]. It is a distributed computing model concentrates on facilitating applications with low latency services. It helps in non-latency aware services. Nearby users use idle computing resources for enhancing the overall performance when all the processing is considered to be extremely higher [14]. These nodes are composed of actuators and sensors. The computation is carried out by the fog devices only in case of emergency conditions and also with the storage facilitates. Similarly, time-sensitive based computation is achieved by the FC process.

B. Fog definition

Researchers define fog computing in various ways [15]. It is explained below:

1. FC is depicted as, “It is highly a virtualized platform that gives storage, compute, and perform services among conventional CC and IoT devices”. However, it is not placed at network edge exclusively.
2. FC is a scenario where an enormous amount of heterogeneous devices like (wireless, ubiquitous, autonomous, and decentralized devices) which potentially perform certain functions and carry out processing tasks and storage without third party interventions. It assists in various services and network functions. The user leases the nodes to host services to get incentives.
3. FC works from a centralized cloud and operates at network ends. It places some resources and processes at end indeed of cloud utilization and storage, respectively.
4. Fog is a distributed computing platform where various non-virtualized and virtualized services carry out the process. Also, it is associated with cloud for performing long-term storage and non-latency aware processing by residing in the middle of the cloud and users.

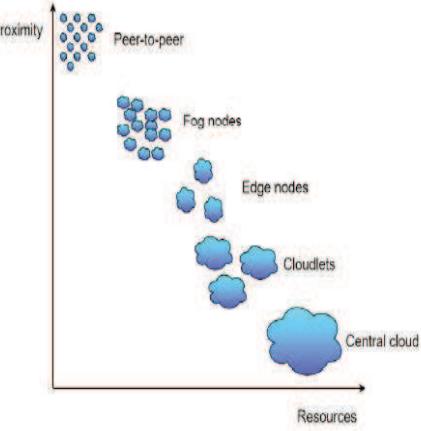


Fig 1: Proximity of fog architecture

C. Fog architecture

It is either application-specific or agnostic. Fig 1 depicts the proximity of fog architecture. The fog moves the application components to some newer locations and initiates earlier processing of certain events. Thus, scalability is achieved. The mobility is determined based on the movement of the nodes. Then, the federation criteria are established based on the coordination execution among the nodes. Finally, interfaces, interoperability are achieved with the data interfaces and control made over the management process.

The fog has the ability to deal with node scaling and dynamic workloads. Thus, the mobility criteria are also fulfilled. Also, it coordinates the execution among the distributed nodes where the federation criteria are not met out. At last, the data interfaces and controls are dealt with the management phase. It can achieve It can achieve interoperability criteria. Fig 2 shows the three-tier architecture of fog.

D. Resource management

When the clouds, fog, IoT are integrated with fog system, the resource management is determined to be a crucial factor. Allocation, scheduling, and resource migration are major research topics to be addressed. Bitten et al., [36] analyzed the resource migration by concentrating on fog nodes and VM migration. The target is to maintain VM when the users move from one location to another. The migration process is carried out when the users do not identify degradation of application performance. The author models layer-based architecture. In the case of IoT device based end-users that the author explains about mobile layers. It is composed of various modules. Application migration module assists in the decision making process during VM migration.

The scalability is achieved with total end-users, fog domains, and fog nodes. Additionally, the author does not concentrate on fog nodes' mobility and mobility criteria. The cooperation is established among the fog and cloud providers. At last, data interfaces and common control are required to facilitate communication among the fog and cloud and to deal with application lifecycle.

The fog broker carries out workload balancing. It has more hosts and selects the more appropriate components based on its utilization, battery, latency, and essential resources are considered for QoS establishment. Here, thousands of components are considered to achieve scalability. The fog system makes use of a specific protocol for communication establishment. The message transferred over this is known as fog messages. It is based on MQTT [24]. It is used for including some meta-data. Moreover, the control interfaces rely on various strata for dealing with the management of application lifecycle. However, interoperability is not achieved over here.

E. Task scheduling

The fog system offers added computational ability towards the edge connected network. In fog, a general query which arises is how to achieve task scheduling. The author, in [25], discusses the task scheduling problem over the network environment with fog layer. Various small layers have the ability to perform computation access among the nodes. The task scheduling strategy works based on two major factors, where the former allocates various computational resources at the individual small cell with various users related to certain objectives. Then, the requests cannot be processed when there is a lack of accessible resources. In the next factor, the clustering is constructed based on the processing. The requests are served based on the order generated. The latency-oriented variants are attained with lower latency and low power consumption. The resource heterogeneity is achieved based on resource availability, and QoS also achieved.

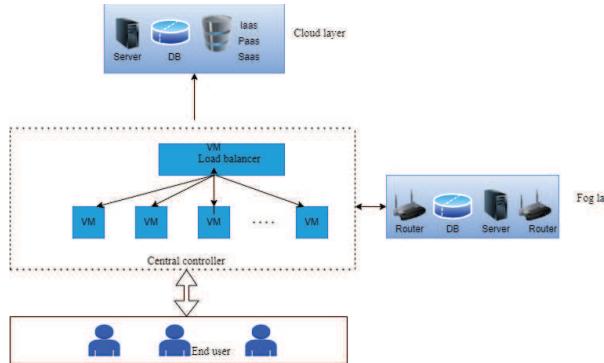


Fig 2: Three-tier fog architecture

Table I: Evaluation criteria limitations and solutions

Evaluation criteria	Research direction	Limitations	Solutions
Heterogeneity	Flexible semantic ontology	The burden in handling heterogeneity	Provide an extension to object-oriented ontology's in cloud system
QoS	SLA management technique	Cloud-based schema extension that pretends to describe the application latency Does not handle all the domain and providers part	Reusability of proper SLA to maintain missing elements like mobile sensitivity, energy-constraints, and resource constraints
Scalability	Adding scaling	Do not provide a general solution to	Integrate and model upstream module with

	resources to the cloud	handle the entire system. Assists either IoT nodes and fog nodes	a global perspective in health checking
Mobility	Mobility of fog and IoT nodes along with the cloud nodes	Concentrates on IoT and fog nodes mobility	Fog assisted mobile entities. Reuse of networking mechanism and adoption of management and placement approaches
Federation	Design of application components	No feasibility and architectural protocol	Distributed composition design module that cooperates and communicates while executing the applications
Interoperability	Signal designing, data interface among the fog systems and domains, control data	No proper inter-domain agreement No operational interfaces for executing the agreements	To define data interfaces, execution of models and control interfaces for data interface.

In the above section, various task scheduling approaches are discussed based on fog where the loads are distributed uniformly over the layers. This provides added stress to concentrate on load balancing and offloading condition over the fog. Hassan et al. concentrate on offloading devices over IoT environment to fog nodes. Fig 3 shows the fog offloading mechanisms.

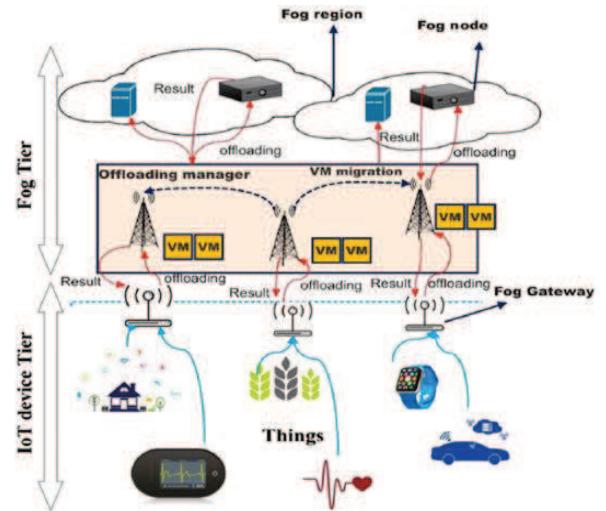


Fig 3: Fog offloading mechanism

The author handles the problems encountered in the mobile devices indeed of offloading. Here, the model interactions are established with the graph structures. The connection is established among the nodes, and interactions are carried among them. The experimental results demonstrate that offloading over

fog layer is efficient than the cloud layer [31]. The offloading process achieves better response time and reduced energy consumption. Li et al., [40] deal with coding schemes and task scheduling over the fog nodes. Here, QoS and heterogeneity criteria are fulfilled. Mobility, federation and elastic scalability are not achieved.

Fricker et al., [42] concentrate on offloading over fog nodes by deploying the data centres over the fog environment. Here, the offloading over the smaller data centres is achieved. When the block request reaches the smaller data centre; then the request is forwarded to the bigger server based on offloading probability.. Here, heterogeneity, QoS criteria, scalability are achieved. However, mobility-based scalability is not achieved. This model operates over the lower-level data centres and does not fulfil the federation criteria.

The algorithm is evaluated based on node removal. The outcomes depict the lower number of moves over the connected edges and rely on the lower migration cost [32]. Therefore, the QoS management is fulfilled. It facilitates the operation at a large-scale environment. Here, scalability and federation criteria are not achieved over here. The mobility criteria are also not provided in a complete form. The author lacks in fulfilling the entire performance metrics. Fig 4 depicts the load distribution among the fog nodes.

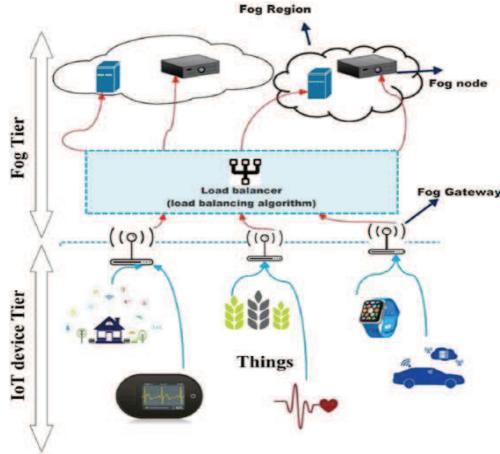


Fig 4: Load distribution among fog nodes

Based on the above analysis, it is observed that load distribution and offloading are concentrated towards the IoT users and fog layers [33]. Various constraints of fulfilled and some are not fulfilled. Various investigators do not achieve some criteria like federation, elastic scalability; mobility scalability.

II. RESEARCH PLATFORM

The fog research platform handles infrastructure and application in fog environment. It deals with resource allocation, scheduling, security, multi-tenancy, fault tolerance and privacy over fog computing. Based on the factors, various research factors are concentrated and discussed below.

a. Task scheduling and resource allocation

Heterogeneous devices are considered as the most challenging factor in the development of proper scheduling and resource allocation. When the nodes want to make use of more computational power when the devices are idle, and the tasks need to be scheduled properly. However, IoT nodes based application in fog faces more complex issues that hinder the latency achievement [34]. The two essential key requirements are task scheduling and resource allocation to achieve efficiency and availability. The resources over the fog nodes are not dedicated, and therefore the resource availability needs to be fulfilled. However, the lack of task scheduling and resource allocation causes unwanted delay during the overall process.

b. Fault tolerance

Fault tolerance needs to be concentrated on during system failure. The failure leads to be network failure, hardware failure, and software failure. The solution results in an operational system where the system pretends to work over the lower capability indeed of shutting the device completely. Generally, fault tolerance is examined over a cloud environment. It is essential to concentrate on fault tolerance in fog. The research work needs to address fault tolerance issues over fog where the investigated actually.

c. Quality of Service (QoS)

In fog, QoS is an essential requirement based on throughput, delay, reliability, and energy consumption. Then, scheduling policies, power consumption, and resource management and power failures need to be addressed to fulfil QoS [34]. When some sensors are failed, the accuracy needs to be affected. Fog intends to work based on latency failure. Moreover, it has to preserve reliability for QoS fulfilment. In some cases, latency criteria need to be addressed. Some author concentrates on various algorithms and methodologies to ensure network connectivity and reliability, accuracy, which acts as a crucial factor for the construction of fog computing environment.

d. Service requirements

The services provided by fog are determined based on the single/multiple requests generated by the users, where the users have to update the service outcome constantly. The services are provided based on the subscription. The outcomes for these requirements are not fixed, and it should be maintained till the service termination. The fog servers and devices need to perform various intermediate processing. This intermediate processing is achieved among service output and user request [35]. The server can communicate with cloud for information retrieval and processing when needed. For example, the selection of best path has to be achieved with real-time traffic over the smart system. The services need to be maintained constantly for fault mitigation, network latency, service quality, power consumption and network latency to preserve various services.

e. Privacy and security

Fog-based computing is distributed completely and not completely managed. The data among the nodes are processed using the intermediate devices where the application does not have complete control over the devices. Users need to transfer data in a secured manner without any harmful entities. The service provides to model security to preserve application from

unnecessary data interruption. There are various types of security factors that need to be fulfilled. They are user privacy, security, and data security. The data and network security are more applicable to both service providers and users perspectives. However, the privacy of users is also essential as the processing is carried out with user data.

f. Multi-tenancy

Methodology	Architecture	Resource allocation	Resource utilization	Resource sharing	Load allocation	Energy efficiency	Load distribution	QoS
Data distribution & resource allocation	Cloud/fog	✓	✓	✓	X	X	X	Fast response time
Generalized decomposition algorithm	Cloud/fog	X	X	X	✓	✓	X	X
Dynamic resource allocation	Fog	✓	X	X	X	X	X	Completion time
Backtracking searching	Cloud/fog	X	X	X	X	X	X	X
Mapping of lower bound algorithm	Cloud/fog	✓	X	X	X	✓	X	Latency
Distributed research sharing	Cloud/fog	X	X	✓	X	X	X	X
Energy-aware algorithm	Fog	✓	X	X	X	✓	✓	Latency
Crowd algorithm	Fog	X	X	✓	X	X	X	X
Dynamic resource computation	Cloud/fog	✓	X	X	X	X	X	X

Process and gives virtualization solution [35]. The container-based virtualization needs to be emulated to facilitate virtualization. Therefore, it is simpler to migrate and manage. It is needed for establishing platform requirements and defined properly. It incurs security issues and performance degradation. Therefore, secure and adequate isolation is essential. Table II explains the limitations of fog computing.

III. FUTURE RESEARCH DIRECTIONS

The following are the future research directions:

1) Resources are more heterogeneous and dynamic due to the diverse nature of devices and resources. These devices are termed as fog devices which is accountable for carrying various computation based on the applications [45]. Fog computation is more dynamic; however, it is predictable through long-term activity analysis. The prediction process is more essential due to fog task execution. When compared to the cloud, it is possible to access more resources and exclusively utilized for cloud requests. Fog targets to use idle resources over the fog devices to perform computation. Therefore, scheduling and resource allocation are more challenging than conventional scheduling and resource allocation.

2) The device failure probability is extremely higher due to the device distribution and management of devices is not so centralized [46]. Therefore, the devices are not failed for various reasons due to software and hardware failure because of a user's activity. Some other factors, like mobility, connectivity, and power sources, have to be concentrated. The devices are connected over wireless connections. It is more obvious that wireless connections are not reliable [47]. The devices are connected wireless and mobile in nature. Therefore, devices can change location to various clusters. The device characteristics are battery-powered, and might leads to failure. Therefore, the

The allocation of multiple tenants for the same service needs to be handled in an isolated manner. It is essential for fog due to the limited constraints in fog environment. By facilitating multi-tenancy, the fog devices run to serve multiple tenants. It is container-based and virtualization enabled. It is a lightweight

Table II: Limitations with fog computing

failure of nature is extremely complex [48]. Therefore, it is essential to fulfilling SLA by evaluating QoS parameters. Also, the fog needs to fulfil uninterrupted connection to fulfil application requirements. The fog devices are accountable for handling the emergency surveillance, and then connectivity failure needs to be handled. When the cloud connectivity is failed, the fog needs to fulfil constant connectivity [49]. Therefore, connectivity among the fog, IoT devices, and cloud should show utmost significance. The connectivity protocol of fog devices is different. Thus, the issues are handled efficiently.

IV. CONCLUSION

Presently, a fog computing paradigm is completely infancy where extensive researches are carried out. In this survey, an extensive analysis is carried out on fog architecture, definition, task scheduling, and resource allocation. Based on this analysis, it is known that heterogeneity, QoS is fulfilled; but lacks in the federation, elastic mobility, and scalability mobility. Finally, there are some open research problems and challenging needs to be addressed. This comprehensive survey provides a strong belief to deal with these issues. The future research directions are also provided in the fast-growing advancements in fog computing. From this analysis, a fog computing paradigm is still an immature research problem.

V. REFERENCE

- [1] G. L. Santos, P. T. Endo, M. F. F. da Silva Lisboa, L. G. F. da Silva, D. Sadok, J. Kelner, T. Lynn, et al. "Analyzing the availability and performance of an e-health system integrated with edge, fog and cloud infrastructures". In Journal of Cloud Computing 7.1 (2018), p. 16.
- [2] C. Yu, B. Lin, P. Guo, W. Zhang, S. Li, and R. He. "Deployment and Dimensioning of Fog Computing-Based Internet of Vehicle Infrastructure for Autonomous Driving".

- In: IEEE Internet of Things Journal 6.1 (2019), pp. 149–160.
- [3] S. W. Loke. “The internet of flying-things: Opportunities and challenges with airborne fog computing and mobile cloud in the clouds”. In: arXiv preprint arXiv:1507.04492 (2015).
- [4] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli. “Fog computing and its role in the internet of things”. In: Proceedings of the first edition of the MCC workshop on Mobile cloud computing. ACM. 2012, pp. 13–16.
- [5] A. Ahmed, H. Arkin, D. Battulga, A. J. Fahs, M. Farhadi, D. Giouroukis, A. Gougeon, F. O. Gutierrez, G. Pierre, P. R. Souza Jr, et al. “Fog Computing Applications: Taxonomy and Requirements”. In arXiv preprint arXiv:1907.11621 (2019).
- [6] R. Mahmud, R. Kotagiri, and R. Buyya. “Fog computing: A taxonomy, survey and future directions”. In: Internet of everything. Springer, 2018, pp. 103–130.
- [7] A. V. Dastjerdi, H. Gupta, R. N. Calheiros, S. K. Ghosh, and R. Buyya. “Fog computing: Principles, architectures, and applications”. In: Internet of things. Elsevier, 2016, pp. 61–75.
- [8] Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu. “Edgecomputing: Vision and challenges”. In: IEEE Internet of things journal 3.5 (2016), pp. 637–646.
- [9] P. Sherubha, “Design of network Threat detection and classification based on Machine learning algorithms in Wireless Sensor Networks”, Caribbean Journal of Science, Volume 53, Issue 2 (May - Aug), 2019.
- [10] Dolci and S. K. Datta. “Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing”. In: 2017 Global Internet of Things Summit (GIoTS). IEEE. 2017, pp. 1–6.
- [11] Satyanarayanan, R. Schuster, M. Ebling, G. Fettweis, H. Flinck, K. Joshi, and K. Sabnani. “An open ecosystem for mobile-cloud convergence”. In: IEEE Communications Magazine 53.3 (2015), pp. 63–70.
- [12] Zhang, Z. Zhang, and H.-C. Chao. “Cooperative fog computing for dealing with big data on the internet of vehicles: Architecture and hierarchical resource management”. In: IEEE Communications Magazine 55.12 (2017), pp. 60–67.
- [13] Agarwal, S. Yadav, and A. K. Yadav. “An efficient architecture and algorithm for resource provisioning in fog computing”. In: International Journal of Information Engineering and Electronic Business 8.1 (2016), p. 48.
- [14] Intharawijit, K. Iida, and H. Koga. “Analysis of fog model considering computing and communication latency in 5G cellular networks”. In: 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops). IEEE. 2016, pp. 1–4.
- [15] Souza, W. Ramirez, X. Masip-Bruin, E. Marin-Tordera, G. Ren, and G. Tashakor. “Handling service allocation in combined fog-cloud scenarios”. In: 2016 IEEE international conference on communications (ICC). IEEE. 2016, pp. 1–5.
- [16] Tang, Z. Chen, G. Hefferman, T. Wei, H. He, and Q. Yang. “A hierarchical distributed fog computing architecture for big data analysis in smart cities”. In: Proceedings of the ASE BigData & Social Informatics 2015. ACM. 2015, p. 28.
- [17] Tuli, R. Mahmud, S. Tuli, and R. Buyya. “FogBus: A Blockchain-based Lightweight Framework for Edge and Fog Computing”. In: Journal of Systems and Software (2019).
- [18] Saurez, K. Hong, D. Lillethun, U. Ramachandran, and B. Ottenwälder. “Incremental deployment and migration of geo-distributed situation awareness applications in the fog”. In: Proceedings of the 10th ACM International Conference on Distributed and Event-based Systems. ACM. 2016, pp. 258–269.
- [19] Chen, Y., Yang, T. Zhang, M.-T. Zhou, X. Luo, and J. K. Zao. “Fog as a service technology”. In: IEEE Communications Magazine 56.11 (2018), pp. 95–101.
- [20] Gill, R. C. Arya, G. S. Wander, and R. Buyya. “Fog-Based Smart Healthcare as a Big Data and Cloud Service for Heart Patients Using IoT”. In: International Conference on Intelligent Data Communication Technologies and the Internet of Things. Springer. 2018, pp. 1376–1383.
- [21] Bittencourt, M. M. Lopes, J. Petri, and O. F. Rana, “Towards Virtual Machine Migration in Fog Computing,” in 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2015, pp. 1–8.
- [22] Cardellini, V. Grassi, F. L. Presti, and M. Nardelli, “On QoS-aware scheduling of data stream applications over fog computing infrastructures,” in 2015 IEEE Symposium on Computers and Communication (ISCC), 2015, pp. 271–276
- [23] Zeng, L. Gu, S. Guo, Z. Cheng, and S. Yu, “Joint Optimization of Task Scheduling and Image Placement in Fog Computing Supported Software-Defined Embedded System,” IEEE Trans. Comput., vol. PP, no. 99, pp. 1–1, 2016.
- [24] Deng, R. Lu, C. Lai, and T. H. Luan, “Towards power consumption-delay trade-off by workload allocation in cloud-fog computing,” in 2015 IEEE International Conference on Communications (ICC), 2015, pp. 3909–3914
- [25] Li, M. A. Maddah-Ali, and A. S. Avestimehr, “Coding for Distributed Fog Computing,” IEEE Commun. Mag., vol. 55, no. 4, pp. 34–40, Apr. 2017.
- [26] Ye, M. Wu, S. Tang, and R. Yu, “Scalable Fog Computing with Service Offloading in Bus Networks,” in 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), 2016, pp. 247–251.
- [27] Fricker, F. Guillemin, P. Robert, and G. Thompson, “Analysis of an Offloading Scheme for Data Centers in the Framework of Fog Computing,” ACM Trans Model Perform Eval Comput Syst, vol. 1, no. 4, p. 16:1–16:18, Sep. 2016.
- [28] Ningning, G. Chao, A. Xingshuo, and Z. Qiang, “Fog computing dynamic load balancing mechanism based on graph repartitioning,” China Commun., vol. 13, no. 3, pp. 156–164, Mar. 2016.
- [29] Ottenwälder, B. Koldehofe, K. Rothermel, and U. Ramachandran, “MigCEP: Operator Migration for Mobility Driven Distributed Complex Event Processing,” in Proceedings of the 7th ACM International Conference on Distributed Event-based Systems, New York, NY, USA, 2013, pp. 183–194.
- [30] S. P. Sasirekha, "Clone Attack Detection Using Random Forest and Multi-Objective Cuckoo Search Classification", IEEE International Conference on Communication and Signal Processing, April 4-6, 2019, India.
- [31] S. P. Sasirekha, P. Sherubha, “An Enhanced Vehicle to Cloud Communication by Prediction Based Machine Learning Approaches”, International conference on computing and Information Technology, 2020, <https://doi.org/10.1109/ICCIT-144147971.2020.9213784>

Performance of Indian Cricket Team in Test Cricket: A comprehensive Data Science analysis

Vishnusai Viswajith Tharoor

*Department of Computer Science and Engineering
Amrita School of Engineering, Coimbatore
Amrita Vishwa Vidyapeetham, India
cb.en.u4cse18068@cb.students.amrita.edu*

Dr Dhanya N.M

*Department of Computer Science and Engineering
Amrita School of Engineering, Coimbatore
Amrita Vishwa Vidyapeetham, India
nm_dhanya@cb.amrita.edu*

Abstract—Cricket is a sport that is widely hailed in India. It is followed fervently in the country, such that some players are worshiped, and are considered holy. Created by shepherds of England, it has now grown significantly, and equipped some of the most advanced technologies to its armory, which help players improve their performances, and also plan strategies against opposition teams. The proposed work aims at analyzing the performance of the Indian Cricket Team in Test format. All oppositions are considered, and the data is analyzed to extract valuable information, which could help serve as an indicator for future performances. Exploratory Data analysis and Visualization was done using python libraries such as pandas, numpy, matplotlib, seaborn. Classification models were implemented for the data based on number of overs, and their results were analyzed. Random Forest Classifier was the most effective among the models implemented, with an accuracy of 75%.

Index Terms—ICT(Indian Cricket Team), Test Cricket, DataFrame, Exploratory Data Analysis(EDA), Machine Learning(ML), Classification.

I. INTRODUCTION

A game of cricket has two teams competing against each other. During the start of the game, a coin toss is done by one of the team captains, and the other team captain makes a call(heads/tails). The winner of the toss makes a decision to either bat or bowl. There are three main entities for a game. They are

- Batter
- Bowler
- Fielder

A batter scores runs, and a bowler takes wickets. The fielder aids the bowler in taking wickets, and also to ensure batters don't get easy runs. A win is determined by the aggregate number of runs scored by the team. An inning is the complete duration of one team's batting/bowling performance. At the end of the match, the team with the greater aggregate of runs, is determined as the winner. However, test match cricket is different. The overall game would comprise of minimum 3 innings, to a maximum of 4 innings in normal circumstances. A team has to play a minimum of one inning, and a maximum of two innings. The game lasts for a total duration of 5 days. Various factors like toss, pitch conditions, weather conditions play a role in determining the winner of the game. The following are the results possible for a given game:

- Win
- Draw (or) Tied
- Loss

The proposed work involves Exploratory Data Analysis on the data-set using python, and derives statistical inferences based on the analysis and visualization results obtained. It also deals with mining rules from the data-set. It also implements classification models on the data-set in order to understand the significance of attributes.

II. RELATED WORK

Various articles related to Test format of cricket were published before the advent of the millennium. However, all exploratory data analysis related to cricket has recently evolved. Comparison of ICT's performance in two periods was done using Association Rule Mining. The first was from 1974 to 2000. The next was from 2001 to 2010. The ICT's performance had substantially improved in the second period considered. [1].The Exploratory data analysis on IPL data , authored by Mohapatra et al., [2] focuses on performing data analysis for IPL data, as well as presenting a model that helps select the team, and also predict the winner for a given IPL game. Regression was mainly used in order to perform the task. It considers various factors of the game like venue, toss, team, track record, and team that is to play [2] .Pandas is the go-to library when it comes to performing data analysis using python. DataFrame is the data structure that is preferred in python. Libraries like seaborn, matplotlib help in visualizing the data and obtain useful results. Bar plots, scatter plots, box plots are some of the plots that could help obtain valuable insights on performances. [3] [6] [8].Classification could be done using supervised and unsupervised methods, depending on the data present. Support Vector Machines are accurate in predictions, and have the ability to handle high dimensions in data. Random Forest method takes in an ensemble of decision trees, and the final class is decided as the major result provided by a large subset of that ensemble. [8] [12]. Random forest method can be performed for both regression and classification. [4] Also, a simple means of classification using an artificial neural network is the Perceptron. It helps identify the linear separability of data as well. [6] [12].

match_id	season	start_date	venue	innings	ball	batting_team	bowling_team	striker	non_striker	...
2603	291352	2007/08	2008-01-02 Sydney Cricket Ground	4	70.2	India	Australia	RP Singh	A Kumble	...
2604	291352	2007/08	2008-01-02 Sydney Cricket Ground	4	70.3	India	Australia	I Sharma	A Kumble	...
2605	291352	2007/08	2008-01-02 Sydney Cricket Ground	4	70.4	India	Australia	I Sharma	A Kumble	...
2606	291352	2007/08	2008-01-02 Sydney Cricket Ground	4	70.5	India	Australia	I Sharma	A Kumble	...

Fig. 1. Sample data - Australia vs India, Sydney-2008

III. DATA SET

Data set is obtained from cricsheet.org . Cricsheet is a collection of projects which collectively provide data for various aspects of cricket. Ball-by-ball raw data folder, pertaining to all test matches from 2004 till January 11,2022 is chosen. The main data set contains information about 627 test matches that have occurred in this period. The raw data has important attributes pertaining to a given delivery, such as start date, batter, bowler, batting team , bowling team no. of overs, run scored of that delivery, whether it resulted in a wicket, and so on. The raw data is available as two types:

- CSV files that comprises ball-by-ball data
- JSON files that comprises both ball by ball data, as well as off-the-field information such as venue, toss winner, winning team, player-of-the-match etc.

ICT has played close to 178 test matches in this period, and all these information are taken for further analysis. Fig. 1 is an illustration of the raw data for a given match

IV. EXPLORATORY DATA ANALYSIS

EDA involves understanding the type of data, their statistical distributions, handling missing data, and analyzing and understanding the data as a whole. [5].

A. Data pre-processing

The raw data has a detailed description of every delivery, for a given match. In order to obtain valuable insights about ICT's performance, pre-processing is necessary. So, for a given test match, off-the-field information such as year, venue, opponent, toss winner, India's role in the first innings, winner, player-of-the-match, were obtained from the JSON file pertaining to that particular match, which was identified using the match-id attribute in Figure 1. The same attribute was used to compute the inning-wise score, wickets taken , and top performance based on runs or wickets from each team. The innings score is computed, when the data is grouped by innings, and the sum of runs scored in that particular inning is computed. A similar approach is followed to compute other continuous attributes. Another attribute present is the Captain for the game. Two captains are considered: MS Dhoni and Virat Kohli. Although other players have captained the team for a few matches, ICT was following major trends during their tenure. This attribute could really help indicate the nature of performances given by the team. Fig. 2 denotes the pre-processed data obtained.

Year	Venue	Opponent	Choice	Captain	Innings	1st innings no. of overs	1st innings score	1st innings wickets lost	1st innings top score	1st innings top wickets	Toss_Winner	Winner	MoM
173	Kempton Oval, London	England	Batting	V Kohli	191	61.3	10	57	4	...	England	India	RG Sharma
174	The Rose Bowl, Southampton	New Zealand	Batting	V Kohli	217	92.1	10	49	5	...	New Zealand	New Zealand	KA Jamieson
175	Wankhede Stadium, Mumbai	New Zealand	Batting	V Kohli	325	109.5	10	150	10	...	India	India	MA Agarwal
176	SuperSport Park, Centurion	South Africa	Batting	V Kohli	327	105.3	10	123	6	...	India	India	KL Rahul
177	The Wanderers Stadium, Johannesburg	South Africa	Batting	V Kohli	202	63.1	10	50	4	...	India	South Africa	D Elgar

Fig. 2. Sample pre-processed data of ICT

B. Data cleaning

After the following data was obtained, data cleaning was performed. On analysis, it was found that the 4th innings mostly had missing values. The reason behind this, is that Test cricket allowed for wins that involved an innnings' defeat. In this case, the winning team doesn't perform an extra inning, as the one inning performed has earned them a win. In some cases matches were drawn due to unfavorable weather or unforeseeable circumstances. Filling the missing value with 0 would necessarily mean that the analysis of the attributes could lead to incorrect findings. Hence, the missing values were replaced with median, as the median is sensitive to outliers, and it also doesn't impact the other values present.

C. Statistical insights

During this period, it was found that the average total by ICT in the 1st innings, is 367, and in the 2nd innings, is 386, and in the 3rd innings ,is 252, and in the 4th innings, is 172. ICT has won 49 games batting first, and 38 games chasing. This indicates that the team has been more successful batting first. Also, the team has drawn 26 games while chasing, and 21 games while batting first. This indicates that it could also prevent a loss by batting sensibly in the last innnings.The most successful chase by the team, in this period, has been 387. The highest total scored by the team, in this period, is 759. The lowest score by the team for any innnings, is 36. In terms of bowling, the best bowler in the team would bag a minimum of 4 wickets, when it came to wins. The no. of times the bowler has taken 10 wickets, is comparatively more in Virat Kohli's leadership, than M.S Dhoni's leadership.

1) *Batters statistics:* Fig.3 denotes the data derived from raw data about batters. The raw data is grouped by innnings, and if the batting team is India, then the details about batters are stored and later converted into a pandas DataFrame. Kohli is the most successful batter during this period. In a span of 11 years, he has scored 7854 runs, with an average of over 50. Conversion rate is the percentage of 50 plus scores that have been converted into hundreds. Kohli has the best conversion rate among batters. VVS Laxman has hit 35 half centuries, the most during the period. MS Dhoni has the worst conversion rate among batters who have played more than 100 innnings. I Sharma has scored the most ducks or zeroes during this period. Strike rate does not play a major role in test cricket. [16].

2) *Bowlers statistics:* Fig. 4 denotes the data derived from raw data about bowlers. The raw data is grouped by innnings, and if the bowling team is India, then the details about bowlers are stored and later converted into a pandas DataFrame.

Player	Innings	Runs	Average	HS	Ducks	200+	100s	50s	Conversion rate
V Kohli	166	7854	51.31	254	14	7	27	27	50.00
CA Pujara	160	6661	45.63	206	11	3	18	32	36.00
SR Tendulkar	133	5969	48.88	248	2	3	17	28	37.78
R Dravid	134	5708	46.60	270	3	1	16	26	38.10
V Sehwag	125	5525	48.20	319	13	3	14	23	37.84

Fig. 3. Best batters of ICT in tests (2004-present)

Ashwin is the most successful bowler in this period. Even though he made his debut in 2011, he has taken the most wickets during this period. The number of five wicket hauls he has taken, is the highest amongst Indian bowlers, by a huge margin. He is also the player who has played the most number of innings as well.

3) *Captaincy statistics:* The data is derived by filtering matches based on wins, draws and losses. Win loss ratio, and win percentage are derived attributes. MS Dhoni and V Kohli have contrasting captaincy styles, and it is reflected in the match results as well. Under MS Dhoni, India had won 31 test matches, and lost 24, which is contrasting the captaincy tenure of V Kohli. India had won 44 games, and lost just 16. In the literacy survey, it was already found that India was better performing in the 21st century, compared to its previous years. This effectively proves that V Kohli is the most successful captain for India in the test format. It is also notable that Virat Kohli has won 7 player-of-the-match awards as captain. Usually, captaincy tends to affect the bowling/batting performance of the player. V Kohli has been the best batter, as well as the best captain for the ICT in the test format, during this period.

4) *Nation wise statistics:* The data is derived, by grouping the pre-processed test matches data based on opposition, and then aggregating the wins, draws and losses based on Dataframe filtering. In this period, ICT has the most wins against Australia and England. It is also notable that in this period, it has played the most number of matches, as well as lost the most number of matches against Australia and England. India has been most successful against the West Indies, as it has played 20 games so far, and hasn't lost a single test match. It can also be concluded that England has been the most difficult team to beat in this period. Overall, India has fared well against all oppositions, as there is no scenario where the number of losses exceed the number of wins.

D. Data Visualization

Data Visualization is a measure that would help users gain insights on data through various graphical plots and illustrations. Cricket analysts use data visualization tools in order to show various useful relations between various data attributes.

Fig. 6 depict ICT's annual aggregate wins in the considered period. This data is derived by filtering the data based on wins achieved by ICT in matches, and aggregating the total

Player	Innings	Wickets	5WH
R Ashwin	140	450	34
I Sharma	138	317	11
RA Jadeja	89	238	9
Harbhajan Singh	87	236	12
Z Khan	90	220	8

Fig. 4. Best bowlers of ICT in tests (2004-present)

Opponent	Total matches	Wins	Draws	Losses
Australia	37	16	10	11
England	40	16	9	15
South Africa	27	13	5	9
West Indies	20	12	8	0
Sri Lanka	19	11	4	4
New Zealand	18	8	6	4
Bangladesh	9	7	2	0
Pakistan	6	2	3	1
Zimbabwe	1	1	0	0
Afghanistan	1	1	0	0

Fig. 5. Performance of ICT against each nation

number for each year. 2016 is the most successful year for ICT , With 9 wins in the test matches held. 2010, 2021 are the second most successful years for Indian Cricket. Additionally, the same steps were repeated to visualize India's toss record, as shown in Fig. 7. India won the most tosses in the year 2008.

The annual number of centuries were derived as a DataFrame from the raw data. Each game in the dataset was grouped by the innings that was being played. Then, the data was filtered based on ICT being the batting team, and the overall sum for each player was aggregated. Year, No. of Centuries are the data involved. 2010 was the most successful year for ICT in terms of the no. of centuries. 23 centuries were hit in that year, signifying the batting was at its peak during that year. 2016 and 2017 saw the most no. of centuries that resulted in wins. Fig. 9 depicts the centuries hit by ICT players.

Fig. 10 denotes that more the number of innings played, more the probability for a batter to score more runs. The trend followed so far, is that if a player exceeds more than 120 innings, the increase in the no. of runs seemed to be following an exponential trend.

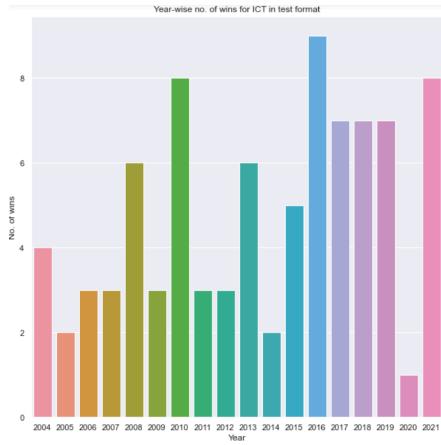


Fig. 6. ICT - Aggregate annual wins

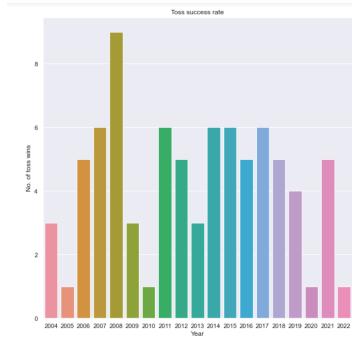


Fig. 7. ICT - Aggregate toss wins

Fig. 8 is a box plot between runs scored in the first innings for each year, based on choice. Blue plots denote the runs conceded, and red plots denote the runs scored. Until 2014, the blue plots mostly were plotted above the reds, denoting that the runs conceded were higher than runs scored. Since 2015, the red plots were consistently plotted higher than the blue plots. Until 2014, MS Dhoni was the Indian Captain. After 2014, V Kohli was the Indian captain. This indicated the improvement of tactics against opponents, and also the improvement in bowling performances.

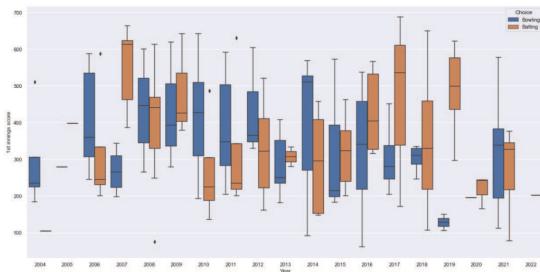


Fig. 8. Box plot between runs scored in 1st innings each year based on choice

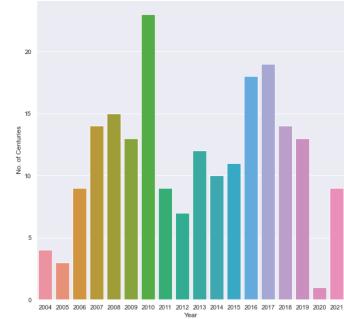


Fig. 9. ICT - Aggregate annual centuries

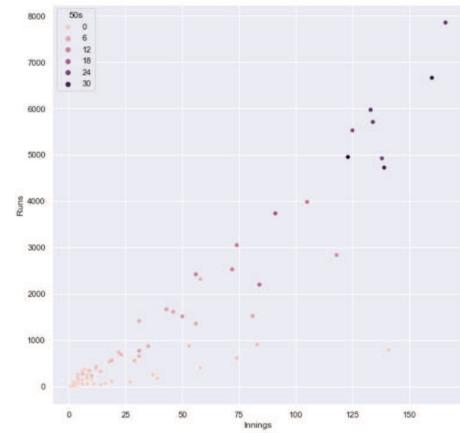


Fig. 10. Scatter plot between innnings and runs scored- ICT

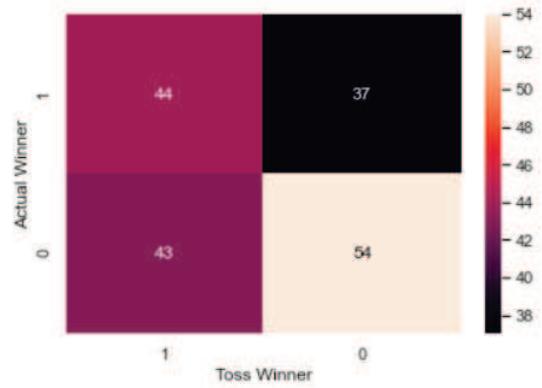


Fig. 11. Heat map of confusion matrix between toss wins and match results

Fig. 11 denotes the heat map of a confusion matrix between toss wins and match results. This was done in order to check the impact of toss result on the overall match results. 55% of the games played, had teams that won the toss, won the match. Unlike one day games, toss plays a vital role in test format, as there may be scenarios where the team that wins the toss would be choosing a decision that would provide suitable

conditions, and thus create winning situations.

Additionally, a bar chart was created to visualize the aggregate player-of-the-match awards given to each player, and from the plot, it was inferred that V Kohli and R Ashwin were awarded the most awards. Both players got 7 awards during this period. It is also notable that V Kohli and R Ashwin are the most successful among batters and bowlers respectively. Thus, the better an individual player performs, the more awards that could be won.

V. ADDITIONAL PRACTICES

A. Association rule mining - Apriori Algorithm

Apriori algorithm is used to find out association rules between objects [1]. Association rules are derived from frequent item-sets, which is a resultant of the algorithm [15]. For the considered data, attributes were considered as objects, and associations between attributes was to be found. The ICT Test matches data was first split into two data sets:

- Data set where ICT batted first
- Data set where ICT bowled first

The first data set was grouped by the 1st innings score, and second was grouped by the 2nd innings score. This was done in order to emphasize on the batting performance on both the innings. After encoding, and ensuring that the data is in the right format, the model is implemented with all proper hyperparameters, and the results are given in Fig. 12 and Fig. 13.

When a deeper understanding of these results are made, it can be understood that when a team plays more no. of overs, the magnitude of the top scorer of the innings would be high. Subsequently, the innings score would be high. This could help set up a good 3rd innings performance and subsequently win the game.

When the team bowls first, the lower the top scorer, the better the bowlers have been able to capitalize on taking wickets. If the 1st innings score is less, then it encourages the team to perform better in the second innings. If a significant score is posted, the team chasing could be easily favourable for a win.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
217	(1st innings no of overs)	(3rd innings score, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
435	(1st innings no of overs, 1st innings score)	(3rd innings score, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
439	(1st innings no of overs)	(1st innings score, 3rd innings top score, 1st inn..)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
543	(1st innings no of overs, Year)	(3rd innings score, 1st innings top score, Year)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
548	(1st innings no of overs)	(3rd innings score, Year, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf

Fig. 12. Apriori algorithm results for batting first

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
299	(1st innings top score)	(2nd innings score, 1st inning score, 3rd inn..)	0.049383	0.491481	0.049383	1.0	2.079923	0.029806	inf
313	(2nd innings no of overs)	(2nd innings score, 1st innings score, 2nd inn..)	0.037037	0.491481	0.037037	1.0	2.079923	0.019204	inf
404	(Year, 1st innings top score)	(3rd innings score, 1st innings score, 2nd inn..)	0.049383	0.491481	0.049383	1.0	2.079923	0.029806	inf
412	(1st innings top score)	(3rd innings score, 1st innings score, 2nd inn..)	0.049383	0.491481	0.049383	1.0	2.079923	0.029806	inf
433	(Year, 2nd innings no of overs)	(3rd innings score, 1st innings score, 2nd inn..)	0.037037	0.491481	0.037037	1.0	2.079923	0.019204	inf

Fig. 13. Apriori algorithm results for bowling first

TABLE I
CLASSIFICATION RESULTS

Classifier model	Accuracy
Random forest classifier	75%
Support Vector Machine	58%

B. Hypothesis Testing

Hypothesis testing is the method used to check if a statistical statement can be accepted or rejected, by means of certain computations [14]. The following tests were done:

- Mean of number of wins each year is greater than 7
- Mean of number of Centuries after 2014 is lesser than 7

These tests are one tailed [14]. Significance value for the hypotheses considered are 0.05 and 0.1 respectively. Both hypotheses were accepted. The first hypothesis is an indicator of how the team has fared better under V Kohli's captaincy. The second hypothesis is an indicator of how the batters are relatively making lesser impact, and how bowlers have gained prominence.

C. Classifier models

Based on the data input, a classifier categorizes it into different classes. In the proposed work, models have been created based on the number of overs. There are two classes: 0 denoting ICT hasn't won the particular game considered; and 1 denoting ICT winning the game. Random Forest Classifier and Support Vector machine were the two models chosen [7]. Random Forest Classifier helps improve the accuracy of the results, due to the consideration of multiple decisions, and Support Vector Machines work very efficiently in linearly separable data. Table I denotes the results obtained. Random Forest Classifiers provided 75% accuracy. The reason for Support vector machines not producing favourable results, may be due to the fact that the data is not linearly separable. It also denotes the fact that the number of overs being played in a test match involving ICT has a role to play in deciding the winner.

VI. CONCLUSION

A. Findings

On performing the study, it can be said that since 2004, there has been an upward progression in terms of the no. of wins, and also in the overall team performance. This has been concluded using the win-loss ratio amongst opponents as the metric to measure the overall success. Various notable players have retired from the game, but it has not affected the overall performance of the team. Since 2014, there has been an upward trend in the overall performance of the team. The number of losses during that period were relatively low, to the losses before this period. Win-loss ratio and ratio between wins and sum of draws and losses, are the metrics used to measure the abilities of captains. It can be inferred through the analysis, that V Kohli has been the best captain in this period. A batter's performance is best measured by the conversion rate. The findings suggest that V Kohli has been the best performer

of the team. Apart from these findings, it can be noted that the overall performance of batters after 2018 has been low, and it needs to be improved. Amongst the classifier models tested, Random Forest Classifier has the best accuracy relatively, to showcase the effect of the no. of overs on the match result.

B. Future enhancements

The data that could be extracted on bowlers were relatively low. More information could be extracted, and analysed. Home-Away match performance could not be interpreted, and based on Geo-Encoding, accurate locations could be extracted, and performance at home, and away, could be analysed. When subsequent matches are played, more data would be thus present. The pandas library has been the go-to solution to handle table-like data in python. Terality is a newly developed library, that has the same syntax as python, but has a relatively better performance in all aspects [10]. This library could be used in order to perform operations. More models, such as deep learning models could be used in order to produce more insights on the data. [11] [16].

REFERENCES

- [1] Raj, K. A., Padma, P. (2013). Application of association rule mining: A case study on team india. 2013 International Conference on Computer Communication and Informatics. <https://doi.org/10.1109/iccci.2013.6466294>
- [2] Mohapatra, S., Goswami, A., Singh, A., Singh, V. K., Sen, B., & Sharma, K. (2021). Exploratory Data Analysis on IPL data. Lecture Notes in Networks and Systems, 315–325. https://doi.org/10.1007/978-981-16-4244-9_26
- [3] Pani, D. S. (2019). Exploratory data analysis using Python. International Journal of Innovative Technology and Exploring Engineering, 8(12), 4727–4735. <https://doi.org/10.35940/ijitee.l3591.1081219>
- [4] Raundale, P., Thosar, C., & Rane, S. (2021). Prediction of parkinson's disease and severity of the disease using machine learning and deep learning algorithm. 2021 2nd International Conference for Emerging Technology (INCET). <https://doi.org/10.1109/incef51464.2021.9456292>
- [5] Saranya, E., & Sivakumar, P. B. (2020). Data-driven prognostics for run-to-failure data employing machine learning models. 2020 International Conference on Inventive Computation Technologies (ICICT). <https://doi.org/10.1109/icict48043.2020.9112411>
- [6] M.S Suchitra., & L Pai, Maya (2017). Machine Learning using Advanced Python. <http://www.csi-india.org/>
- [7] Sivaranjani, S., Ananya, S., Aravindh, J., Karthika, R. (2021). Diabetes prediction using machine learning algorithms with feature selection and dimensionality reduction. 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). <https://doi.org/10.1109/icaccs51430.2021.9441935>
- [8] Pisner, D. A., & Schnyer, D. M. (2020). Support Vector Machine. Machine Learning, 101–121. <https://doi.org/10.1016/b978-0-12-815739-8.00006-7>
- [9] Pajankar, A. (2022). Hands-on Matplotlib. <https://doi.org/10.1007/978-1-4842-7410-1>
- [10] <https://towardsdatascience.com/good-bye-pandas-meet-terality-its-evil-twin-with-identical-syntax-455b42f33a6d>
- [11] <https://www.geeksforgeeks.org/ipl-score-prediction-using-deep-learning/>
- [12] Dogru, N., &; Subasi, A. (2018). Traffic accident detection using random forest classifier. 2018 15th Learning and Technology Conference (L&T). <https://doi.org/10.1109/ltn.2018.8368509>
- [13] <https://www.geeksforgeeks.org/implementing-apriori-algorithm-in-python/>
- [14] <https://towardsdatascience.com/hypothesis-testing-in-machine-learning-using-python-a0dc89e169ce>
- [15] <https://www.geeksforgeeks.org/apriori-algorithm/>
- [16] Sarkar S, Mishra S, Kumar S. Development of a Comprehensive Multi-Factor Method for Comparing Batting Performances in One-Day International Cricket. IIM Kozhikode Society & Management Review. 2022;11(1):92-108. doi:10.1177/22779752211017260

DESIGN OF DEEP LEARNING BASED COVID 19 DIAGNOSIS

FRAMEWORK USING LUNG ULTRASOUND IMAGES

Dr Rajesh Doss
Assistant Professor, NDA, Pune -23
extrajesh@gmail.com

ABSTRACT

The global health catastrophe caused by the Coronavirus disease pandemic (COVID-19) and related control efforts has impacted every aspect of human life. The most important requirement for COVID-19 diagnosis is early detection of the condition. The ML algorithm aids in the acceleration of the process while also conserving energy. Time-to-delivery and the availability of training data, on the other hand, are critical. Deep learning algorithms surpass covid 19-based lung ultrasound scans in diagnosing them, according to a thorough background analysis. As a result, this study shows how to develop a CNN-based framework for Lung Ultrasound indicators in COVID-19 in real time. This research looks into the roadmap of lung ultrasonography indicators in detail, with a focus on COVID-19. Finally, this article emphasizes the investigation of the covid19 problem in different domains.

Keywords:

Coronavirus 2, COVID-19, Deep Learning, Lung Ultrasound, Machine Learning.

1. INTRODUCTION

Lung infections that are potentially fatal Coronavirus 2 is the third most prevalent human

pathogenic coronavirus, capable of causing pneumonia in 15-20% of infected people and necessitating emergency treatment in 5-10% of cases. It all began in Wuhan, China, and has since spread throughout the world. Other signs of a serious illness include pneumonia, lymphopenia, lymphocyte depletion, and cytokine release syndrome. As the COVID-19 virus spreads, quick and reliable testing and decision-making tools are in high demand. Clinical trials are required for diagnosis because the external signs are identical to those of the flu [4].

Prompt diagnosis and touch monitoring are important components of COVID-19 emergency preparedness in order to prevent the virus from spreading further. Health care providers can examine the impact of critical management decisions as a result of the influx of new patients, particularly those in need of emergency care. Although CT is a well-established method of detecting COVID-19, it has several limitations: it is not widely available, reaction times are long, and patients must be moved out of their unit via procedure. It's nearly impossible to use CT equipment properly during a pandemic, and doing so can quickly exhaust scarce resources [8][11].

One of the most significant obstacles to the COVID-19 inquiry is a lack of precise and sufficient

data. Multiple death reports and virus-affected diseases go uninvestigated due to a lack of testing. It's hard to say whether the COVID-19 infection detection failure factor is three, 300, or even more. No country in the world has been able to produce credible information on this topic. Research and development, on the other hand, must continue, necessitating knowledge fusion [14].

Millions of people all around the world have been affected by the current pandemic. Hundreds of thousands of individuals have been afflicted with this highly contagious disease, raising concerns about humanity's long-term sustainability. Only by catching the sickness early and avoiding infecting others can it be kept under control. This necessitates a precise and rapid diagnosis that poses no health risks. Traditional machine learning algorithms struggle to provide the same results due to obstacles such as detection time, cleaning requirements after each use of the diagnostic apparatus, and resource availability. As a result, this study created a deep learning-based CNN-based classifier for COVID diagnosis using Lung Ultrasound indicators. Section II examines the linked study, Section III discusses the CNN design for LUS markers and the findings, Section IV examines the covid19 problem in various locations, and Section V finishes with a conclusion.

2. RELATED STUDY

COVID-19 has prompted the creation of a wide range of unique medical technologies, from telemedicine to remote sensing, as a result of the global epidemic. At the same time, the epidemic is taxing the healthcare system. Medical imaging, which

includes everything from chest radiography to computed tomography and thoracic ultrasound, will be used to identify and treat Coronavirus infections. The use of Artificial Intelligence (AI) in the COVID-19 management MI that has been implemented so far in this demanding testing, as well as clinical application, was studied. According to this study, physicians and AI groups place different emphasis on imaging modalities and performed tasks [3][12].

Lung ultrasonography (LUS) has recently acquired popularity for diagnosing COVID-19 pneumonia. Several articles on its application based on hypothesis research, case reports, or a series of retrospective cases have yet to be published, as well as the predictive status of LUS in COVID-19 patients. LUS was employed in a case-control study to see if it could predict death and ICU admission in COVID-19 patients who were screened in the ER. According to our findings, at the initial assessment in the emergency department, LUS might diagnose COVID-19 pneumonia and predict high-risk patients for ICU admission and mortality [8][10].

The fastest-growing segment of ultrasound technology is lung ultrasonography (LUS). Several doctors throughout the world used LUS during the current COVID-19 outbreak to diagnose lung illness in patients who were suspected or infected with the virus. Standard ultrasound imaging, on the other hand, is commonly used to create LUS, despite the fact that it is not designed to withstand the high pressures encountered in lung tissue. The LUS, as well as the computing algorithms, are not yet configured. Certain aspects of LUS demand scientific

inquiry and treatment in order to enhance the usage of alternative ultrasound imaging. This overview attempts to give you a clear picture of what happened [1][5].

In LUS images with B-lines of various etiologies, a convolutional neural network was fixed. CNN's diagnosis accuracy was comparable to that of other LUS doctors utilizing a 10% data recovery set. Two Canadian tertiary hospitals are used as natural facilities. People can't detect the difference between the advent of LUS pathology, such as COVID-19, and an in-depth study model can. The fact that biomarkers are not visible in existing ultrasound images and that multidisciplinary investigations have been proven is demonstrated by the difference in picture quality between humans and models [2][6].

To assess the consistency of in-depth research methods of the COVID-19 diagnostic process, the LUS database of patients with COVID-19, bacterial pneumonia, non-COVID-19 non-COVID-19 pneumonia, and eligible participants was constructed. We believe that a neural-based convolutional neural network with high sensitivity and specificity may appropriately identify the COVID-19 LUS.

3. ANALYSIS OF COVID19 CRISIS IN LUS

3.1 LUS

If active infections are recognised early, they can be prevented and controlled. Patients with minor illnesses do not need to be admitted to the hospital unless they are in grave danger of deteriorating. In the short run, a comprehensive plan to aid health care providers in identifying patients and evaluating the possibility of developing a critical or critical state, or

shifting from critical to violent scenarios, can help hospitals better access restricted resources. Exiled patients can be closely followed at home with LUS imaging. This is especially important in settings with high-quality hospital beds, such as long-term care facilities. If there aren't enough COVID-19 test kits, LUS can aid diagnose patients[7].

LUS is a non-invasive supplemental screening method that can be used in any situation. Patients who are low-risk and those who are high-risk could be distinguished using a preliminary screening. Medical experts can easily do LUS in the case of sickness. For patients who must be admitted to the hospital, this will also allow for more precise pre-hospital treatment. By capturing LUS images next to the bed, the number of health providers who can encounter the patient should be minimized. The use of a chest X-ray or CT scan now necessitates referring the patient to a radiology institution, potentially exposing additional people to the infection[9].

3.2 Comparison analysis of classification methods

Table 1. Comparison analysis of classification methods

Method	Accuracy (%)	Sensitivity(%)	Specificity(%)
ANN (Barrientos et al.[2016])	-	91.5	100
CNN (Kulhare et al.[2018])	92	96	79
SVM (Carrer et al.[2020])	94	-	-

RVM Sangeetha et al.[2021])	100	100	100
Supervised Feed Forward (Correa et al.[2018])	-	90.9	100
Stochastic approach (Brattain et al.[2013])	100	-	-

LUS pictures are classified using one of two methods: 1) Featured segregation, in which individual pieces are analyzed stochastically, and 2) read-based techniques, such as NNs, which act as a black box solution. Stochastic methods are a straightforward, straightforward approach that may be better suited to a portable LUS system. Classified classifiers analyze classed images and categorize image information using statistical retranslation and image filtering. RVMs, unlike SVMs, identify potential rather than ambiguous detection. Ultrasound image accuracy is a critical component of performance. This means that the imaging tool should be coupled with a portable ultrasound system and should advise health care personnel during the LUS exam to ensure proper image quality independent of technology or operator[13][15]. In [17] disparity measure is computed based on the frequency of dimensional match where the divergence is estimated from. Estimated according to the disease support measures using the data's value for disease support enhances the accuracy of the classification.

4. SYSTEM METHODOLOGY

Wireless transducers and tablets are excellent ultrasonic equipment for diagnostics. The first step in

the detection process is to increase image quality, which can then be used to eliminate unwanted image detail, also known as image noise[16]. In raw picture photos, there exist distortions. This preprocessing is essential due to the poor contrast of frames in terms of inconsistencies, irregular borders, and objects. A number of filters can be used to minimize noise.

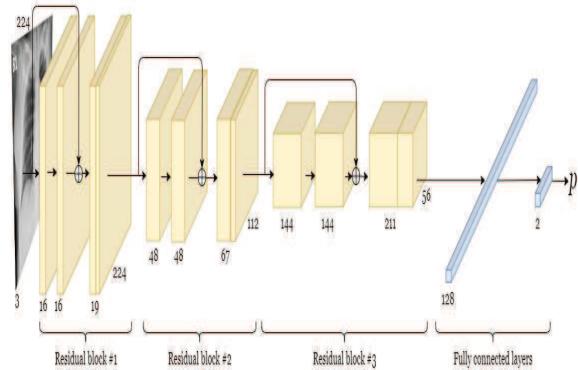


Figure 1. Convolutional neural network

The pictures from Covid 19 are converted to grayscale and then smoothed. A segmentation is the process of separating an area of interest (ROI) from the frame sense. The method is based on a supervised learning model that makes use of the complex learning properties of neural network models. Enough samples of covid infected pictures are used for training to transfer knowledge to the neural network model. The number of features and their convergence towards the optimal segmentation solution are determined by the optimization method, whose objective function is defined as.

$$S_K(x, y) = O_{min}(F(x, y)) \in R_n; n = 1, 2, 3 \dots 255$$

where

$S_k(x, y)$ denotes the segmented output,

O_{min} defines the objective minimization function and $F(x, y)$ denoting the feature vectors reflecting the pre-processed covid image to be segmented.

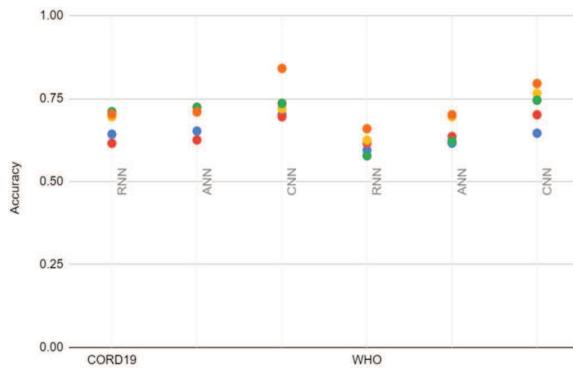


Figure 2. Accuracy analysis

The efficiency and reduction in computation time were evaluated against the numbers of images in the dataset ranging from 10 to 100 and in each case the proposed CNN model outperforms traditional learning based models supporting the simplicity of the proposed solution, which at the same time shows improvement in accuracy.



Figure 3. Computation analysis

5. CONCLUSION

Pneumonia caused by COVID-19 Dialysis patients can be treated straight away with LUS tests

performed at the bedside, which is typical in dialysis units with limited treatment alternatives. Ultrasonography may be employed in patients with COVID-19, despite the fact that few articles support it. We discuss our experiences with point-of-care ultrasonography, notably lung ultrasound, as well as the advantages of in-depth learning methodologies in this study. The work also shows how CNN is designed and how it outperforms RNN and ANN in terms of accuracy.

REFERENCES

- Arntfield, Robert & Vanberlo, Blake & Alaifan, Thamer & Phelps, Nathan & White, Matt & Chaudhary, Rushil & Ho, Jordan & Wu, Derek. (2020). Development of a deep learning classifier to accurately distinguish COVID-19 from look-a-like pathology on lung ultrasound. [10.1101/2020.10.13.20212258](https://doi.org/10.1101/2020.10.13.20212258).
- Barrientos R., Roman-Gonzalez A., Barrientos F., Solis L., Correa M., Pajuelo M., et al. (2016). Automatic detection of pneumonia analyzing ultrasound digital images. 2016 IEEE 36th Central American and Panama Convention (CONCAPAN XXXVI), San Jose, CA, November 9–11, 2016, (New York, NY: IEEE;), 1–4.
- Bonadì, Nicola & Carnicelli, Annamaria & Piano, Alfonso & Buonsenso, Danilo & Emanuele, Gilardi & Kadhim, Cristina & Torelli, Enrico & Petrucci, Martina & Maurizio, Luca & Biasucci, Daniele & Fuorlo, Mariella & Forte, Evelina & Zaccaria, Raffaella & Franceschi, Francesco. (2020). Lung Ultrasound Findings Are Associated with Mortality and Need for Intensive Care Admission in COVID-19 Patients Evaluated in the Emergency Department. Ultrasound in Medicine & Biology. 46. [10.1016/j.ultrasmedbio.2020.07.005](https://doi.org/10.1016/j.ultrasmedbio.2020.07.005).
- Born, Jannis & Beymer, David & Rajan, Deepa & Coy, Adam & Mukherjee, Vandana & Manica, Matteo & Prasanna, Prasanth & Ballah, Deddeh & Guindy, Michal & Shaham, Dorith & Shah, Pallav & Karteris, Emmanouil & Robertus, Jan & Gabrani, Maria & Rosen-Zvi, Michal. (2020). On the Role of Artificial Intelligence in Medical Imaging of COVID-19. [10.1101/2020.09.02.20187096](https://doi.org/10.1101/2020.09.02.20187096).
- Born, Jannis & Brändle, Gabriel & Cossio, Manuel & Disdier, Marion & Goulet, Julie & Roulin, Jérémie & Wiedemann, Nina. (2020). POCOVID-Net: Automatic Detection of COVID-19 From a New Lung Ultrasound Imaging Dataset (POCUS).

6. Born, Jannis & Weidemann, N & Cossio, M & Buhre, Charlotte & Brändle, Gabriel & Leidermann, K & Aujayeb, Avinash & Rieck, Bastian & Borgwardt, K. (2021). L2 Accelerating COVID-19 differential diagnosis with explainable ultrasound image analysis: an AI tool. *Thorax*. 76. A230.2-A231. 10.1136/thorax-2020-BTSabstracts.404.
7. Brattain L. J., Telfer B. A., Liteplo A. S., Noble V. E. (2013). Automated b-line scoring on thoracic sonography. *J. Ultrasound Med.* 32, 2185–2190. 10.7863/ultra.32.12.2185
8. Carrer L, Donini E, Marinelli D, Zanetti M, Mento F, Torri E, Smargiassi A, Inchingolo R, Soldati G, Demi L, Bovolo F, Bruzzone L. Automatic pleural line extraction and covid-19 scoring from lung ultrasound data. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2020 Nov; 67(11):2207-2217.
9. Correa M, Zimic M, Barrientos F, Barrientos R, Román-Gonzalez A, Pajuelo MJ, Anticona C, Mayta H, Alva A, Solis-Vasquez L, Figueroa DA, Chavez MA, Lavarello R, Castañeda B, Paz-Soldán VA, Checkley W, Gilman RH, Oberhelman R. Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition. *PLoS One*. 2018; 13(12):e0206410.
10. Cristiana, Balloescu & Grzegorz, Toporek & Seungsoo, Kim & Katelyn, McNamara & Rachel, Liu & Shaw, Melissa & Mcnamara, Robert & Balasundar, Raju & Moore, Christopher. (2020). Automated Lung Ultrasound B-Line Assessment Using a Deep Learning Algorithm. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. PP. 1-1. 10.1109/TUFFC.2020.3002249.
11. Demi L. Lung ultrasound: The future ahead and the lessons learned from COVID-19. *J Acoust Soc Am*. 2020 Oct;148(4):2146. doi: 10.1121/10.0002183. PMID: 33138522; PMCID: PMC7857508.
12. Desai, Sudhen & Pareek, Anuj & Lungren, Matthew. (2020). Deep learning and its role in COVID-19 medical imaging. *Intelligence-Based Medicine*. 3-4. 100013. 10.1016/j.ibmed.2020.100013.
13. Kulhare S., Zheng X., Mehanian C., Gregory C., Zhu M., Gregory K., et al. (2018). Ultrasound-based detection of lung abnormalities using single shot detection convolutional neural networks. *Simulation, image processing, and ultrasound systems for assisted diagnosis and navigation*. (Springer), 65–73.
14. Roy, Subhankar & Menapace, Willi & Oei, Sebastiaan & Luijten, Ben & Fini, Enrico & Saltori, Cristiano & Huijben, Iris & Chennakeshava, Nishith & Mento, Federico & Sentelli, Alessandro & Peschiera, Emanuele & Trevisan, Riccardo & Maschietto, Giovanni & Torri, Elena & Inchingolo, Riccardo & Smargiassi, Andrea & Soldati, Gino & Rota, Paolo & Passerini, Andrea &
- Demi, Libertario. (2020). Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. *IEEE Transactions on Medical Imaging*. PP. 10.1109/TMI.2020.2994459.
15. Smargiassi, Andrea & Soldati, Gino & Torri, Elena & Mento, Federico & Milardi, Domenico & Giacomo, Paola & De Matteis, Giuseppe & Burzo, Maria & Larici, Anna & Pompili, Maurizio & Demi, Libertario & Inchingolo, Riccardo. (2020). Lung Ultrasound for COVID-19 Patchy Pneumonia: Extended or Limited Evaluations?. *Journal of ultrasound in medicine: official journal of the American Institute of Ultrasound in Medicine*. 40. 10.1002/jum.15428.
16. S.K.B.Sangeetha, R.Dhaya, Dhruv T Shah,R.Dharanidharan, and K. Praneeth Sai Reddy, "An Empirical Analysis of Machine Learning Frameworks Digital Pathology in Medical Science", *Journal of Physics: Conference Series*, 1767, 012031, doi:10.1088/1742-6596/1767/1/012031,2021.
17. Ananthajothi.K, Subramaniam.M , 'CLDC: Efficient Classification of Medical Data Using Class Level Disease Convergence Divergence Measure', *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN 2278-3075, Vol. 8, no. 10, pp. 2256-2262 (2019).<https://doi:10.35940/ijitee.J1123.0881019>

Artificial Intelligence Model for Classification of Sensitivity Data Utilized at Industrial Applications

Ashraf Mohammad*
Department of Electronic Engineering
National Ilan University
Yilan City, Taiwan
mdashraf211199@gmail.com
(corresponding author)

Muddineni Raveendra
Department of Electronic Engineering
National Ilan University
Yilan City, Taiwan
muddineniravi962@gmail.com

Mupparaju Kavya Sree
Department of CSE
Vel Tech Rangarajan Dr Sagunthala
R&D Institute of Science and
Technology
Chennai, India
mupparajukavyasree@gmail.com

Yaswanth Pavan Kumar. T
Failure Data Analysis Engineer
Wistron corporation
Hsinchu City,Taiwan
yaswanthpavan8@gmail.com

Saam Prasanth Dheeraj Pedapalli
Department of ECE
Vel Tech Rangarajan Dr Sagunthala
R&D Institute of Science and
Technology
Chennai, India
dheeraj.saam@gmail.com

Abstract— Every electronic gadget in the structure has the potential to cause ignite, whether through a short circuit or another means which is also the primary cause of fire accidents. Smoke detectors are used to save the lives of causalities during hazards. In the manufacturing industry of smoke detectors, around 12 percent of products are getting failed due to the parameters of sensitivity criteria. This paper provides an approach to improve the success rate sensitivity data of the products as an industrial application using a sequential model provided by the tensor flow. Model accuracy and loss were results compared at different epochs. The final artificial intelligence model was developed with a minimum number of layers and requires less computational time. This paper helps the industry to improve product quality and efficiency.

Keywords— sequential model, sensitivity criteria, tensor flow, product quality, computational time

I. INTRODUCTION

An apparent dispersion of carbon or other airborne particles released by a burning substance has been alluded to as smoke. It is typically an unpleasant metabolic end of fires and fireplaces, but it may also be utilized for pest extermination. The leading cause of mortality among victims of indoor fires is smoke inhalation. Many of the chemicals in fire smoke are very toxic and poisonous. Carbon monoxide is the most harmful, resulting in carbon monoxide poisoning. Smokers who inhale a substantial amount of smoke rapidly get incapacitated and lose consciousness. Furthermore, a plume of smoke that comes into touch with the ambient atmosphere has the potential to be ignited by another ignition source in the vicinity or by its own warmth. This results in phenomena such as back draught and short circuits [1]. To detect the presence of smoke, a smoke detection sensor is employed. Every electronic device in the building has the potential to ignite, whether by a short circuit or another mechanism. If the detection system is deployed in the appropriate position, many lives can be saved. A low-cost safety equipment is a fire detection system. So there's no excuse for not having a fire detection system in your home or work place.

The product is transported to the manufacturing line for parameter updates, modifications, and a series of testing after it has been produced and circuit tested in the factory. As part of the parameter setup. During the parameter updating process, the factory places the circuit board on a dedicated machine, allowing the machine to calculate the required

parameters based on the circuit board and the circuit characteristics, then write the parameters into the circuit board, and forwarded it to the parameter adjustment machine according to the specifications to do the parameter adjustment in the way of testing, resulting in the product's adjusted parameters. However, these pre-adjusted settings may often cause issues beyond their own. After entering the adjustment, the failure rate is extremely high, or they will definitely fail, but they may not be screened out and re-entered the first time, but they must actually enter the test adjustment process only subsequently could an issue be found as well as returned to the production line, even if the number of times is excessive, manual adjustment is required, which greatly affects production efficiency.

AI deep learning technology is an indispensable technology for smart factories. AI deep learning technology belongs to the use of data application level. There are several application directions in smart factories, using deep learning technology to predict whether device data If there is an abnormality, is there any place to be adjusted or replaced; image detection can also be used to analyse the product quality, which can increase efficiency and reduce labour costs; it can also be used to analyse the product's market situation to help the decision-making team adjust the production line. Introducing AI technology into Industry 4.0 is the key point that academia, industry and government agencies must cooperate to research and make breakthroughs.

This paper is organized as follows previous research works and the goal of the project is discussed in Section II. The project structure and methodology were described in Section III. Results of several experiments which were made on the proposed methodology were discussed in Section IV and finally conclusion in Section V.

II. BACKGROUND WORK

To begin, make sure to know about the four architectural levels of the smart factory. The physical resource layer refers to the ability to sense components, interconnection, data collection, etc. between manufacturing systems, in order to record various activities in the factory. data application layer use data analysis and scientific decision-making methods to plan production plans, control product quality, adjust and calibrate equipment, the network layer can use the network to store the data that originally existed in the local server into the cloud server at the final terminal layer through people and

computer network inside the organization, to establish a global cooperation process for smart manufacturing in the order-driven market. The fundamental goal of the smart factory is to enhance quality management, lower labour expenses, increase production efficiency, and so on, in order to achieve a level beyond human management's capabilities. This is the future smart factory's strategic orientation.

The fourth industrial revolution (Industry 4.0) usually refers to production automation, intelligence, and networking. Because of the emergence of a large number of new technologies, including big data [2], Internet of Things (IoT) [3] so as early as 2013, the German government proposed a technology plan called Industry 4.0, also known as the fourth industrial revolution, and Industry 4.0 is closely related to the Industrial Internet of Things (IIoT) [4], and the specific practice is the concept of the smart factory. In [5] several data preprocessing papers was discussed to maintain and explore massive datasets.

In this work, factory currently generates a large amount of data throughout manufacturing. As mentioned in the previous section, the majority of these data are sensitivity adjustment parameters. However, once generated, the majority of this data are stored in the factory's server and cannot be used routinely. Because of the enormous amount of data, humans may not be able to identify a relationship between these parameters. But, there may be many correlations between these data that affect the success or failure of product adjustment. A new methodology was proposed in the following section to overcome this kind of issue in the manufacturing line.

III. METHODOLOGY

In this section, we have described the techniques which we have used to develop an artificial intelligence model from the raw data format which was obtained from the manufacturing industry for the purpose of the project.

A. Data Pre-processing

Various data will be incorporated in terms of data integration. Incorrect data to verify the rationality and completeness of the data; inconsistent data to avoid inconsistencies in the data before and after the integration, such as Unit conversion and the same data in different fields; duplicate data to pay attention to the duplication problem that will occur when the data table is merged; There is a recognized mathematical relation between the data in redundancy.

In terms of data cleaning, consider removing missing values from the data using direct deletion, default values, mean values by statistical methods; then use the bounding method and data adaptation method to remove noise problems from the data; and finally, deal with the outlier problem using cluster analysis, regular drawing by removing from the dataset.

The data characteristics are scaled to a certain interval by normalization, and the data is normalized using the average value and standard deviation of a through normalization procedure.

Finally, using the data-by-data reduction procedure, remove any data that is unrepresentative or worthless. In a summary, the objective is to simplify the cleaned and integrated data even further in order to achieve the following goals: enhance the quality of exploration, reduce analysis time, and generate shorter algorithms.

B. Model Selection

Keras models may be built in two ways either sequentially or functionally. For most issues, the sequential API allows you to build models layer by layer. It is restricted however it does not allow you to design models with numerous inputs or outputs or that share layers. Conversely, the functionality API enables you to design models with much greater freedom, since you can easily define models in which levels link to more than only the previous and next layers as shown in figure 1. In reality, layering can be literally connected to another layer. As a result, a sophisticated network such as residue systems may be created. multilayer Perceptron model was defined for classification.

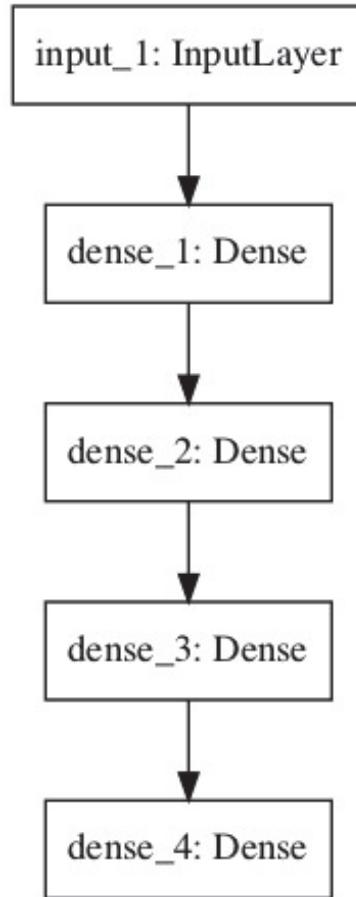


Fig 1. Flow chart of keras sequential model

C. ADAM Optimisation

Adaptive Moment Estimation (Adam) was based on adaptive estimations of lower-order moments that compute adaptive learning rates for each parameter., Adam is a first-order gradient-based optimization technique for stochastic optimal solutions. The approach is simple to develop, computationally intensive, requires minimal storage, is insensitive to gradients scale parameter, which is well adapted for issues with a lot of data and its parameters.

Similarly, to momentum, Adam stores an enormously declining average of past gradients m_t . Although momentum can be compared to a ball rolling down a hill, Adam acts more like a gyro ball with friction, preferring flat minima on the error function. As shown below, we estimate the decay

Fig 2. Original data gathered from the factory server.

Fig 3. Decrypted data.

averages of past and past gradient descent, m_t and v_t separately.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (1)$$

$$v_t = \beta_1 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2)$$

m_t and v_t are the estimations of both the gradients' first moment that is mean and the second moment that is variance. The researchers of Adam notice that because m_t and v_t are introduced as vectors of 0's, they are skewed towards zero, especially during the early time steps and while the degradation rates are low that is β_1 and β_2 are near to one.

The researchers recommend the following default parameters i.e., $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-8}$ [6]. They demonstrate empirically that Adam performs well in practice and measures up towards other adaptive learning-method algorithms.

IV. EXPERIMENTATION & RESULTS

Original data which was collected from the factory has many parameters so data pre-processing is required to remove all the null values and make the data cleaner. Parameters required for labelling the data are considered and the remaining are popped from the sheet. All the parameters' headers are changed to make the data more feasible and convenient to analyze. As shown in figure 2 some of the columns are encrypted as per the factory privacy policy. Data was decrypted and the original information was extracted from these columns.

There are many null values in the data after the extraction of the encrypted data as shown in figure 3. The sequential model was selected and the final dataset as shown in figure 4 which was prepared by selecting the required parameters based on the relationship which affects the output label that is success rate of the product. The remaining parameters Gain, Time, Current, LTD, Norm are the product

parameters initialized by the factory these values were trained and their results are shown in the following figures 5-8. Model loss was gradually reduced when model got trained at each epoch. Subsequently model accuracy was increased. This model comparison with different epochs was shown in figure 9.

In training data, if the label represents 1 it means that particular product has failed the sensitivity test so the parameters have to be re-initialized the label was shown in figure 4.

	A	B	C	D	E	F
1	Gain	Time	Current	LTD	Norm	Label
2	3	200	100	18	22	0
3	2	300	100	15	20	0
4	2	300	100	19	26	0
5	2	400	100	18	23	0
6	1	400	100	13	17	1
7	1	400	100	17	22	0
8	2	400	100	18	22	1
9	2	300	100	16	22	0
10	2	300	100	16	21	1
11	2	300	100	17	22	0
12	2	300	100	17	22	0

Fig 4. Final training dataset.

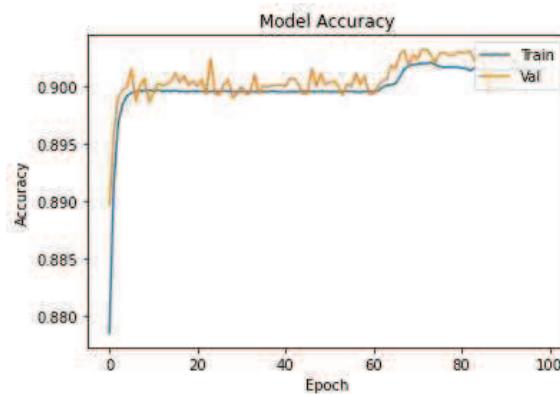


Fig 5(a). Result of model accuracy at 100 epochs.

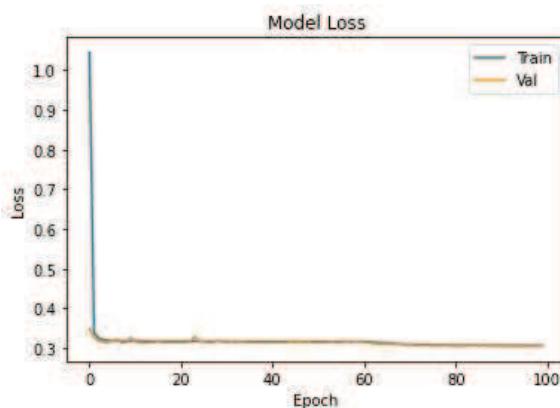


Fig 5(b). Result of model loss at 100 epochs.

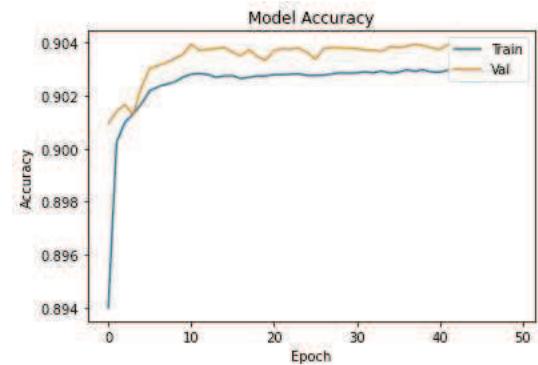


Fig 6(a). Result of model accuracy at 50 epochs.

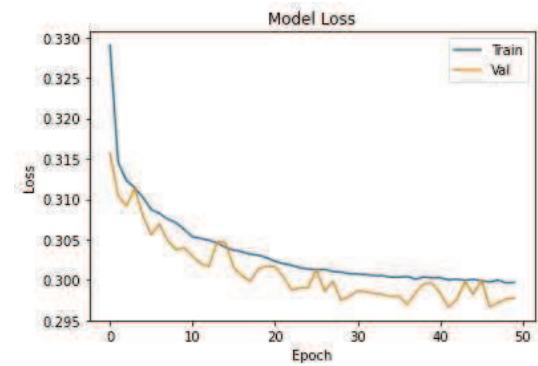


Fig 6(b). Result of model loss at 50 epochs

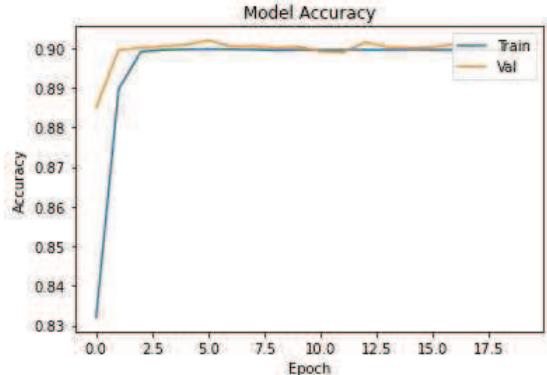


Fig 7(a). Result of model accuracy at 20 epochs.

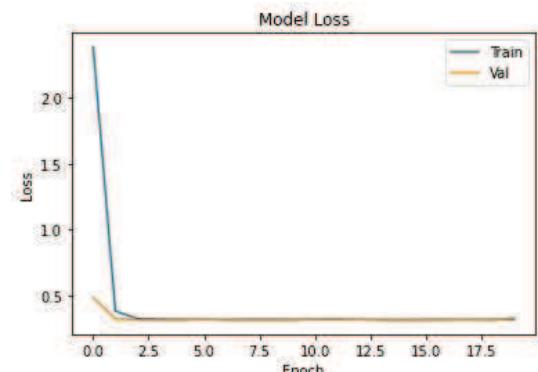


Fig 7(b). Result of model loss at 20 epochs.

```

Epoch 1/20
1/1791 [=====] - ETA: 0s - loss: 256.8948 - acc: 0.1211WARNING:tensorflow:Callbacks method
`on_train_batch_end` is slow compared to the batch time (batch time: 0.000s vs `on_train_batch_end` time: 0.0010s). Check your
callbacks.
1791/1791 [=====] - 1s 555us/step - loss: 6.1058 - acc: 0.8620 - val_loss: 0.3194 - val_acc: 0.9024
Epoch 2/20
1791/1791 [=====] - 1s 489us/step - loss: 0.3246 - acc: 0.9005 - val_loss: 0.3264 - val_acc: 0.9023
Epoch 3/20
1791/1791 [=====] - 1s 494us/step - loss: 0.3266 - acc: 0.9000 - val_loss: 0.3154 - val_acc: 0.9004
Epoch 4/20
1791/1791 [=====] - 1s 495us/step - loss: 0.3252 - acc: 0.8997 - val_loss: 0.3184 - val_acc: 0.9014
Epoch 5/20
1791/1791 [=====] - 1s 485us/step - loss: 0.3256 - acc: 0.8997 - val_loss: 0.3472 - val_acc: 0.8984
Epoch 6/20
1791/1791 [=====] - 1s 502us/step - loss: 0.3273 - acc: 0.8994 - val_loss: 0.3179 - val_acc: 0.9012
Epoch 7/20
1791/1791 [=====] - 1s 489us/step - loss: 0.3252 - acc: 0.8995 - val_loss: 0.3151 - val_acc: 0.9006
Epoch 8/20
1791/1791 [=====] - 1s 516us/step - loss: 0.3259 - acc: 0.8993 - val_loss: 0.3221 - val_acc: 0.9021
Epoch 9/20
1791/1791 [=====] - 1s 516us/step - loss: 0.3251 - acc: 0.8995 - val_loss: 0.3147 - val_acc: 0.9008
Epoch 10/20
1791/1791 [=====] - 1s 486us/step - loss: 0.3247 - acc: 0.8993 - val_loss: 0.3145 - val_acc: 0.9003
Epoch 11/20
1791/1791 [=====] - 1s 492us/step - loss: 0.3260 - acc: 0.8993 - val_loss: 0.3335 - val_acc: 0.8988
Epoch 12/20
1791/1791 [=====] - 1s 490us/step - loss: 0.3242 - acc: 0.8994 - val_loss: 0.3382 - val_acc: 0.8986
Epoch 13/20
1791/1791 [=====] - 1s 484us/step - loss: 0.3248 - acc: 0.8994 - val_loss: 0.3198 - val_acc: 0.9008
Epoch 14/20
1791/1791 [=====] - 1s 489us/step - loss: 0.3248 - acc: 0.8993 - val_loss: 0.3144 - val_acc: 0.9000
Epoch 15/20
1791/1791 [=====] - 1s 486us/step - loss: 0.3242 - acc: 0.8995 - val_loss: 0.3845 - val_acc: 0.8965
Epoch 16/20
1791/1791 [=====] - 1s 486us/step - loss: 0.3253 - acc: 0.8994 - val_loss: 0.3193 - val_acc: 0.8993
Epoch 17/20
1791/1791 [=====] - 1s 485us/step - loss: 0.3248 - acc: 0.8994 - val_loss: 0.3211 - val_acc: 0.8992
Epoch 18/20
1791/1791 [=====] - 1s 485us/step - loss: 0.3248 - acc: 0.8993 - val_loss: 0.3182 - val_acc: 0.8999
Epoch 19/20
1791/1791 [=====] - 1s 489us/step - loss: 0.3244 - acc: 0.8994 - val_loss: 0.3179 - val_acc: 0.9020
Epoch 20/20
1791/1791 [=====] - 1s 486us/step - loss: 0.3233 - acc: 0.8994 - val_loss: 0.3263 - val_acc: 0.8989
Model: "sequential"

Layer (type)          Output Shape         Param #
dense (Dense)        (None, 8)           48
dense_1 (Dense)       (None, 2)           18
Total params: 66
Trainable params: 66
Non-trainable params: 0

```

Fig 8. Model structure.

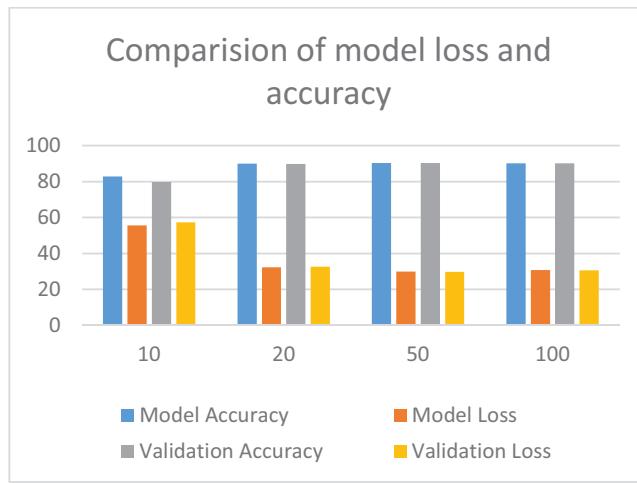


Fig 9. Comparison chart for model loss and model accuracy.

V. CONCLUSION

An Artificial Intelligence model was proposed with an optimized layer. Dataset which was used throughout this project was taken from the alarm manufacturing industry. Sensitivity data was decrypted from the industry server data. Data pre-processing techniques were performed to remove the null values and make the data clear. Model structure was developed by training data and finally, it achieved model

accuracy of 89.94% with model loss of 32.33% and at the validation achieved model accuracy of 89.89% with loss of 32.63%. Furthermore, optimization techniques can be used to reduce the loss by increasing the accuracy model with less computation time. This paper helps the industry to make defected alarm in the production line.

REFERENCES

- [1] G. Xu, Q. Zhang, D. Liu, G. Lin, J. Wang and Y. Zhang, "Adversarial Adaptation From Synthesis to Reality in Fast Detector for Smoke Detection," in IEEE Access, vol. 7, pp. 29471-29483, 2019, doi: 10.1109/ACCESS.2019.2902606.
- [2] G. Xu, Q. Zhang, D. Liu, G. Lin, J. Wang and Y. Zhang, "Adversarial Adaptation From Synthesis to Reality in Fast Detector for Smoke Detection," in IEEE Access, vol. 7, pp. 29471-29483, 2019, doi: 10.1109/ACCESS.2019.2902606.
- [3] M. Younan, M. Elhoseny, A. E.-M. A. Ali and E. H. Houssein, "Data Reduction Model for Balancing Indexing and Securing Resources in the Internet-of-Things Applications," in IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5953-5972, 1 April, 2021, doi: 10.1109/IOT.2020.3035248.
- [4] K. -K. R. Choo, S. Gritzalis and J. H. Park, "Cryptographic Solutions for Industrial Internet-of-Things: Research Challenges and Opportunities," in IEEE Transactions on Industrial Informatics, vol. 14, no. 8, pp. 3567-3569, Aug. 2018, doi: 10.1109/TII.2018.2841049.
- [5] A. Alvarez-Ayllon, M. Palomo-Duarte and J. -M. Dodero, "Interactive Data Exploration of Distributed Raw Files: A Systematic Mapping Study," in IEEE Access, vol. 7, pp. 10691-10717, 2019, doi: 10.1109/ACCESS.2018.2882244.
- [6] Yi, D., Ahn, J., & Ji, S. (2020). An Effective Optimization Method for Machine Learning Based on ADAM. Applied Sciences, 10(3), 1073. <https://doi.org/10.3390/app10031073>

Smart V2G/G2V Charging for Grid connected-Electric Vehicles in Indian Scenario

¹Setlem Bhargavi, ²Uppuluri Baby Jyothsna, ³K.Deepa, ⁴V.Sailaja, ⁵J.Ramprabhakar

Department of Electrical and Electronics Engineering,

Amrita School of Engineering, Bengaluru, India

Amrita Vishwa Vidhyapeetham, India

bhargavisetlem@gmail.com, joshichowdary1830@gmail.com, k_deepa@blr.amrita.edu, v_sailaja@blr.amrita.edu, j_ramprabhakar@blr.amrita.edu

Abstract— In India, most of the vehicles which are used for transportation are fueled by fossil fuels which result in environmental pollution. To reduce pollution, the Government of India has introduced electric vehicles as an alternate. This paper presents smart charging for future Indian grid-connected electric vehicles within the demand response (DR) and regulation specified limits. However, due to few requirements in the Indian electricity market, individual EVs are required to be cumulated by EV charging service providers (CSPs). These CSPs optimizes the costs of charging the EV by scheduling them at the proper time considering electricity rates. This method optimizes the scheduling of Electric vehicles by charging only at the time of off-peak loads (Grid-to-Vehicle (G2V) charging operations of EV) and giving back to the grid at time on-peak loads (Vehicle-to-Grid (V2G) discharging operations of EV). The parameters like time of arrival, departure, initial state of charge (SOC) and the capacity of an Electric vehicle are taken into consideration. Residential load profile, industrial loads profile, total power from different sources like solar, hydro, and Time-of-Use (TOU) tariffs for India is considered for this study. For this Indian data case study, the effect of V2G/G2V of eighty EV profiles is presented and analyzed. An effective algorithm is also proposed and the results show the superiority of the optimized and scheduled operation of EVs by the proposed method.

Keywords— Distribution system, Demand response, Regulation market, Optimal scheduling, Charging service Provider

I. INTRODUCTION

The demand for automobiles is increasing in India. These vehicles run on diesel, gasoline, or natural gas, all of which are important sources of carbon emissions that contribute to global warming and pollution. Additionally, they contribute to the rapid depletion of fossil fuels. The most immediate solution is to figure out how to operate motorized vehicles. Due to environmental concerns, the Indian government has recently boosted the production of electric vehicles, which are more flexible and reliable than motor vehicles [1-3].

Electricity, on the other hand, makes all residential appliances work globally. In addition to all domestic and industrial applications, the transportation sector is now entirely reliant on electricity, with the added benefit of zero-emissions when compared to fuel-based vehicles. Plug-in Electric Vehicle (PEV) and Plug-in Hybrid Electric Vehicle (PHEV) technologies are in high demand for lowering air pollution and greenhouse gas emissions. Plug-in electric cars (PEV) and Plug-in Hybrid Electric Automobiles (PHEV) are vehicles that charge from an external source of power, such as the electric grid (PHEV). In India, the transportation sector would have replaced 30% of its vehicles with electric

vehicles, necessitating the development of ultra-fast charging technologies capable of charging a large number of EVs in a matter of minutes. Although this rapid and efficient charging technology presents several problems for vehicle owners and CSPs, it also puts the distribution power at risk during peak loads (demand load). Charging a high number of EVs at the same time from the grid produces a slew of issues in the power system due to insufficient grid power during peak demands. As a result, a massive analysis of the impact of charging EVs from the grid is required. This improved transportation system necessitates consuming less electricity at peak loads (i.e., returning power to the grid). Electric vehicle charging with Renewable Energy Sources (RES) results in zero pollution and EVs are only charged during off-peak demand load periods under smart charging from grids. These EVs also have the advantage of being able to contribute back to the grid by discharging their power during on-peak demand load periods, accepting all requirements, and alleviating grid issues [4-6].

The methodologies for EV scheduling vary slightly from one another, with V2G operation being used to schedule charging periods. EVs may operate in both grid-to-vehicle (G2V) and vehicle-to-grid (V2G) modes, allowing power to flow in both directions and providing a feasible solution for the various issues that the traditional grid has as a result of higher peak load situations. Hence, EVs are variable loads that may be regulated to provide grid regulation services. EVs can also provide highly desirable flexibility in the power system by utilizing the batteries of parked EVs if a communication infrastructure is in place. For the charging service provider (CSP), these skills provide potential gains from the electricity market. For EVs to participate in power markets, demand response and regulation are important parameters to be taken care. Although individual EVs are restricted from participating in the market due to minimum resource size constraints, aggregation technology allows EVs to provide services to the electric grid. EV owners can make extra income to cover charging costs if they give regulation services or demand reduction to the electric grid. How to identify the appropriate capacity for regulating service and demand response is a fundamental challenge in smart charge scheduling. Although a lot of studies suggest that scheduling set points can help this problem. This paper tries to deploy a smart charging algorithm to face the future charging challenges in an Indian scenario if the EV loads are introduced to the current power level produced. Section 2 deals with the smart charging for charging service providers, section 3 deals with the Indian scenario of power produced, consumed and a case study is presented in section4,[6-10].

II. SMART CHARGING FOR CHARGING SERVICE PROVIDER

Smart charging mandates the participation of CSPs in the Electricity market by providing demand response and regulation. Fig.1. shows the overall participation of CSP in the power market. During the scheduling period, the CSP calculates the hourly set point of individual EVs for charging. The charging rate of the EVs is allowed to vary around the setpoint, once the setpoints have been defined within the given regulation capability, the CSP can provide regulation to the electric grid. In addition, demand response in the energy market can be achieved by modifying the set point of electric vehicles in response to the changes in electricity prices.

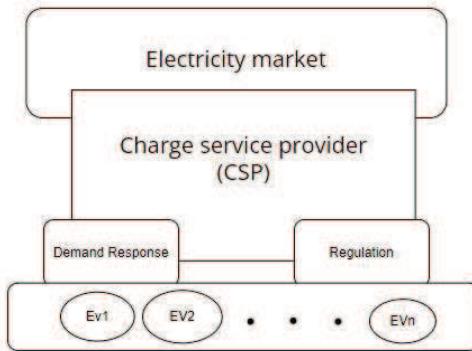


Fig.1. Overview of CSP market participation for charging EVs

CSP's main objective is to charge the EV battery in the time specified by the owner. To optimize profit, it is vital to charge EVs during the cheapest times feasible to finish charging EVs on time. Furthermore, CSPs might profit more from the electricity market, and also should assess their available capacity to bid in the regulation and DR markets when they participate in the electrical markets, considering the varied variables of each vehicle, such as departure, arrival time, the initial state of charge level, and battery power capacity.

III. INDIAN POWER AND LOAD SCENARIO

The charging requirement of EVs depends on the Initial State of Charge (SOC) of the battery, power demand, starting time of arrival, and departure of each EV. The analysis is carried out for an Indian scenario for both industrial and residential types of loads and solar and hydropower produced in a day. The objective of the optimization problem is to determine the setpoints, regulation capacity, and demand reduction capacity of each EV [10-13].

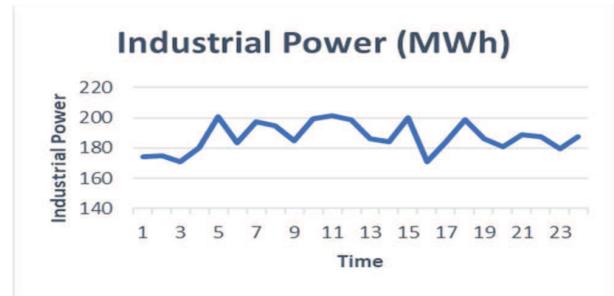


Fig.2. Industrial load

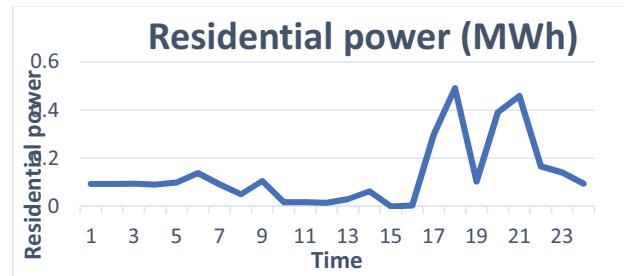


Fig.3. Residential load

Fig.2. and Fig.3. shows the usage of industrial Power and residential Power has taken for the whole single day having time on the x-axis and power on the y-axis, Industrial power is almost ranging from 150Mw -200Mw, Residential power is a peak at 17hrs to 19 hrs and 21hrs to 23 hrs, this helps in further calculating the setpoint.CSPs fix the EVs as a load /source depending on the grid demand, hence to decide, SOC calculations are a must and are detailed below equations.

a) SOC and Capacity can be calculated by:

$$Ahr = \frac{capacity}{voltage} \quad (1)$$

$$SOC = \left(\frac{ISOC}{100} \right) * voltage * Ahr * \frac{100}{1000} \quad (2)$$

$$capacity = capacity(16KW) - SOC \quad (3)$$

Where is capacity the battery capacity of EV, ISOC is the initial state of charge of the EV battery, SOC is the state of charge of a vehicle at time t capacity.

b) Setpoint can be calculated by:

$$SP = Tg - Tu \quad (4)$$

Where Tg is total power generated, Tu is total power used.

c) Cost of CSP to charge the EV batteries:

$$Cost = \sum_{t=1}^n p_t \cdot SP \quad (5)$$

where p_t is the charging price at time t; and SP represents the set point after demand reduction or the expected charging power of EV_i at time t.

d) Benefits of CSP in electricity markets:

$$Income_{it} = \sum_{i=1}^{NV} (g_r \cdot reg_{i,t} + g_{dr} \cdot DR_{i,t}) \quad (6)$$

$$\text{Maximize} = \sum_{t=1}^n \text{income} - \text{cost} \quad (7)$$

Where g_r, g_{dr} is regulation and Demand response at time t, respectively; $reg_{i,t}, DR_{i,t}$ is regulation and Demand response capacities allocated to each EV_i at time t respectively.

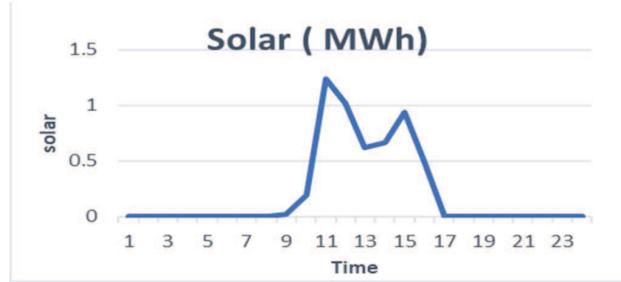


Fig.4. Solar power produced in a day

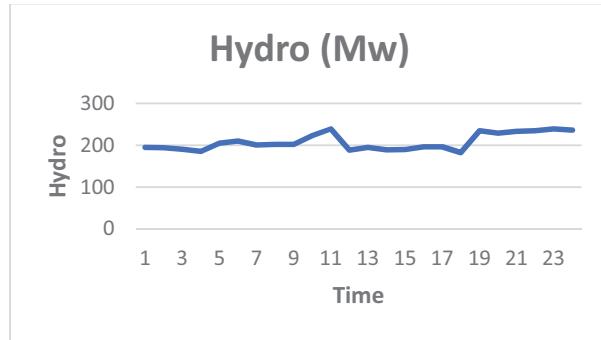


Fig.5. Hydropower produced in a day



Fig.6. Total power generated

Fig.4. shows total solar power generated from an Indian solar firm with a maximum peak at 12hr and 15hr , optimal power from 9 hr to 12 hr, 13hr to 15hr. Fig.5. shows the hydropower produced at an average of 200MW per day. Fig.6. shows the total generated power for both day and the night time having a maximum peak at morning hours nearly at 9 am -10 am and rest with 200MWh level of power[13-15].

Fig.7. shows the algorithm of electric vehicle charging using regulation and demand response for the EV to be scheduled for charging using CSP. Firstly CSP acquires the EV data like ISOC, arrival and departure time, and calculates SOC and capacity, so that it get transparency of EVs charging time, and it has the data of all loads also. When the setpoint is negative at any time it indicates that grid needs power from CSP, hence CSP tries to give back the power to the grid, in return at high selling rate CSP manages to charge EV only at the time of off-peak load . When set point is positive it will charge the EV and if (arrival time-present time) >1 then it will go for departure else it is used for the power usage if any setpoint decrements happen in further hours,. In the context of a multi-agent problem-solving framework, a Mixed Integer-Linear Programming (MILP) approach is provided that allows optimal makespans to be determined for complex categories of scheduling issues that require several steps different aspects into consideration, so that it reduces the charging cost of each EV which makes it more economical to any EV owner.

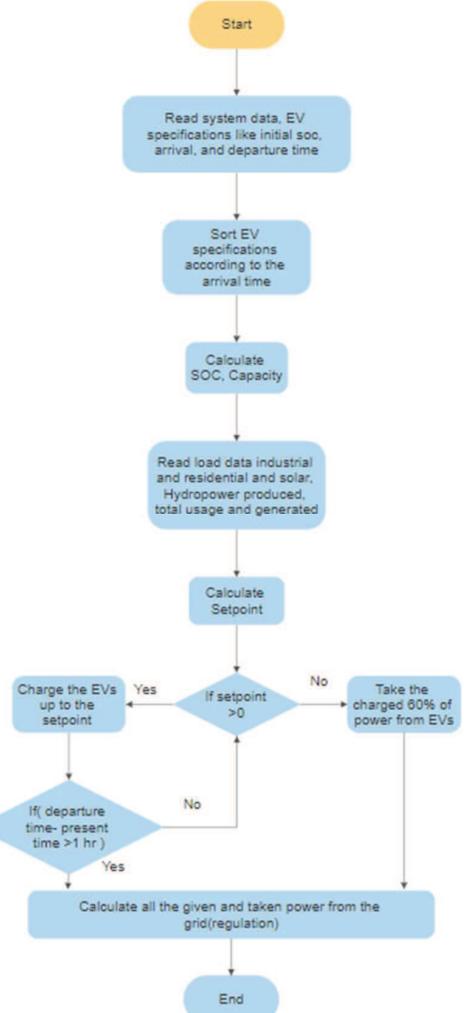


Fig.7. Flow chart of charging EVs using CSP

IV. CASE STUDY

The proposed method is tested for the case study undertaken with 80 private vehicles. The battery capacity (cap) of each EV is considered to be 16KWh, 220V. The market price of regulation capacity (g_r) is considered to be Rs.85/KW as a flat rate. It is also assumed that the DR market is only open from 14h to 16h, and its cost (g_{dr}) is equal to be Rs.150/KW. The setpoint to charge all the batteries and regulations are calculated using the graph shown in Fig.8a. for Karnataka tariff rate in India with the cost price for 1 kW of power at Rs.60, selling price for 1 KW of power at Rs.85, and demand response of Rs.150.

The proposed method was tested for a test case of 80 EV taken under study of the initial SoC, arrival and departure time is shown in Fig.8b., Fig.8c. shows the result of TOU for day and nighttime scheduling that is being used for both residential and commercial purpose. Fig.9. shows the result of the charging set point for the day and nighttime scheduling. In the daily schedule, the EV charging demand has a maximum peak in the morning and evening due to the lower TOU rate and driving departure time. In the night schedule, the EV charging demand peaked in the evening due to the driving departure time, but the charging demand increases at 1h due to the lower TOU rate during the midnight and the negative charging setpoint indicates the power need for the grid hence CSP acquires 60% of charge from the charged EVs and again recharges them in the night hours or the off-peak demand load[10-13].

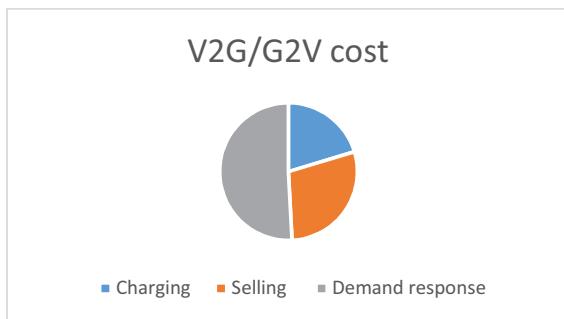


Fig.8a. V2G/G2V cost profile

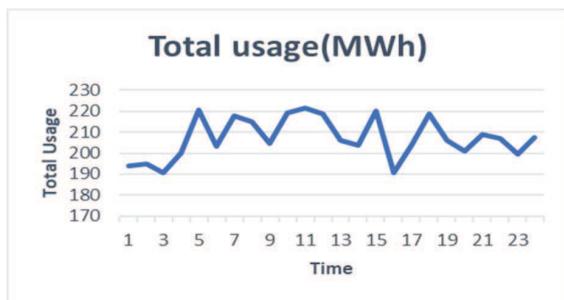


Fig.8b. Power usage

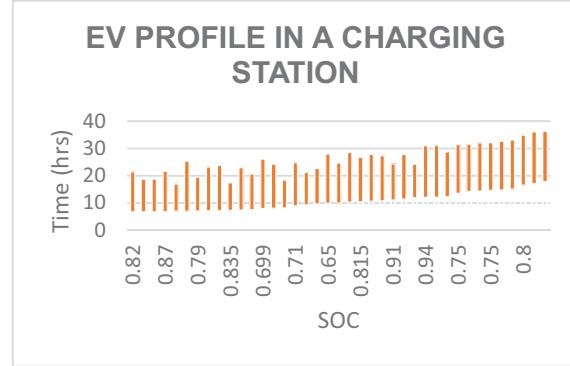


Fig.8c. V2G/G2V cost profile

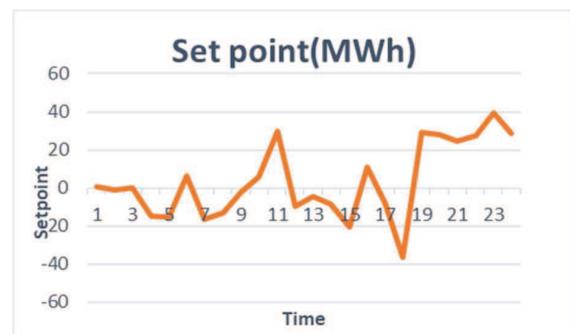


Fig.9. Charging Setpoint



Fig.10. Regulation

Fig.10. shows the results of total regulation capacity and demand reduction for day and night time schedule. During daytime scheduling, many EVs are reducing their charging powers due to the high charging rate. The regulation capacity peaked at 10hr and 5hr and it has a medium peak at 16hr, 18hr to 19hr. This is mainly dominated by higher charging rates at a certain point in time. During night-time scheduling, the regulation capacity is mainly dominated by EV charging power because the EV regulation capacity is limited under its charging power.

Fig.11. shows the result of income and cost while charging EVs with regulation and DR markets and the profit obtained from it. The charging cost is being reduced because the CSPs are giving back the power to the grid at the time of peak load it profits and charging costs are reduced. From equations profit of Rs.75.37464 has been calculated for 1MWpower.

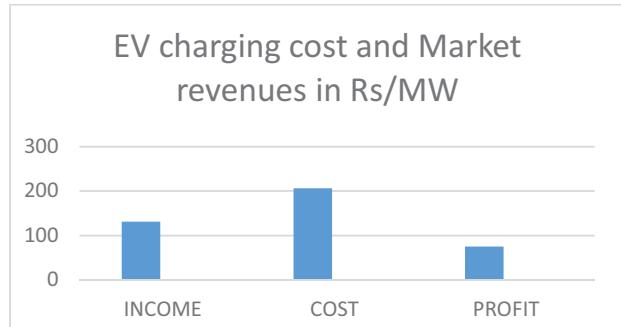


Fig.11. EVs charging costs and market revenues

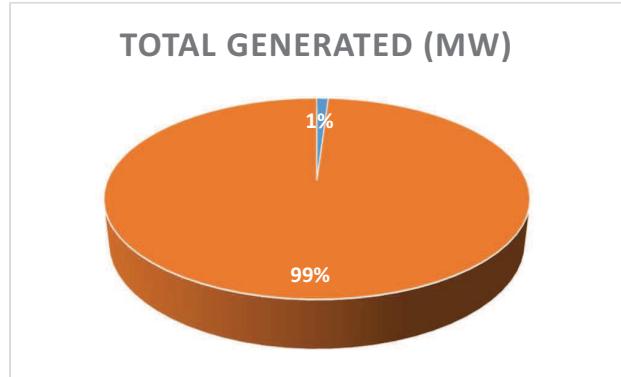


Fig.12. Solar and hydro powers generated.

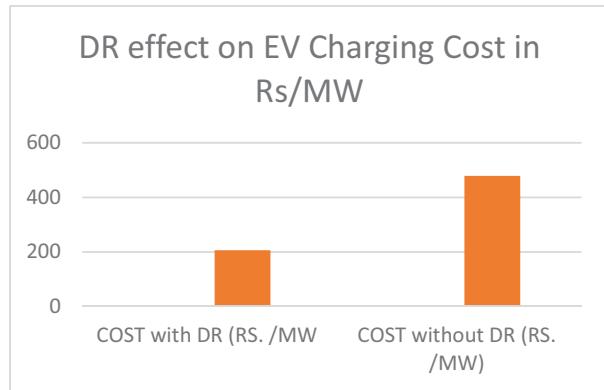


Fig.13. DR effect on EV charging cost.

Fig.12 shows the 1% of solar energy added to the hydro energy generated and the dependency of EVs on the on grid power, and the load on the grid can be minimized if more solar power is pumped into the grid. If solar power can be increased to 10% it can help the grid more efficiently at the time of on-peak demand loads. The effect of Demand response is also studied in Fig.13. which specifies the impact of DR and without the DR. The EVs charging cost using DR was Rs.206.34 which is less when compared to EVs charging cost without using DR (Rs.478.58). Thus DR reduces the charging cost by 50%[17].

V. CONCLUSION

This paper implemented an optimized scheduling method for smart Electric Vehicle V2G/G2V charging to maximize the benefits of a CSP from the DR market as well as to minimize charging costs for an Indian scenario considered. Numerical simulation results with 80 EVs considering both the load profiles, i.e., residential as well as commercial with different tariff schemes and sources like solar and Hydro for Indian scenario are analyzed. The charging setpoint is calculated from the tariff and load profiles, considering that setpoint, EVs are charged. At the time of on-demand loads, the EV's 60% charge power is given back by the CSP to the grid to support the grid. Selling Cost is high so that the CSP is at profit, income is also increased by 25% because of regulation and Demand response. With appropriate scheduling of EV charging, more economic benefits can be expected. Future work will involve methodologies for unexpected departures of EVs with stochastic scheduling. EV is going to be advantageous in terms of air pollution, global warming. It serves as an energy reservoir and provides profit to the CSPs is also a viable solution for the Indian grid to handle a large number of EVs charging.

REFERENCES

- [1] A Sharma, MA Kapoor, S Chakrabarti " Impact of plug-in Electric vehicles on power distribution systems of major cities of India": case study, 2019.
- [2] A. Kapoor, P. Gangwar, A. Sharma and A. Mohapatra, "Multi-Objective Framework for Optimal Scheduling of Electric Vehicles," 2020 21st National Power Systems Conference (NPSC), 2020.
- [3] D. Thiruvonasundari and K. Deepa, "Optimized Passive Cell Balancing for Fast Charging in Electric Vehicle," IETE Journal of Research, 2021.
- [4] T. Duraisamy and Dr. K. Deepa, "Electric Vehicle Battery Modelling Methods Based on the State of Charge-Review", Journal of Green Engineering, 2020.
- [5] R. George, S. Vaidyanathan, and K. Deepa, "Ev Charging Station Locator With Slot Booking System," 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC), 2019.
- [6] K. Poornesh, K. P. Nivya and K. Sireesha, "A Comparative study on Electric Vehicle and Internal Combustion Engine Vehicles," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020.
- [7] B. S. T. Reddy, K. S. Reddy, K. Deepa and K. Sireesha, "A FLC based Automated CC-CV Charging through SEPIC for EV using Fuel Cell," International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2020.
- [8] B. Prasanth, Deepa, K., Jeyanthi R., and B. S., "An Optimized Regenerative Braking System for Electrical Vehicles", IGI Global , 2021.
- [9] K. KUMAR, S. GUPTA and S. NEMA, "A review of dynamic charging of electric vehicles," 2021 7th International Conference on Electrical Energy Systems (ICEES), 2021.
- [10] R. Gauthami, Nair, V. V., Sathish, A., K. Soureesh, V., Ilango Karuppasamy, Sreelekshmi, R. S., Ilangoan, S. A., and Sujatha, S., "Design and Implementation of Efficient Energy Management System in Electric Vehicles", Innovations in Electrical and Electronics Engineering. Springer Singapore, Singapore, 2020.
- [11] Y. Zhiyong et al., "Optimization of Orderly Charge and Discharge Scheduling of Electric Vehicles and Photovoltaic in Industrial Par," 2020 Asia Energy and Electrical Engineering Symposium (AEEES), 2020.
- [12] Agrawal, Manas & Raja Patel, Mohammad Sufiyan, "Global Perspective on Electric Vehicle", International Journal of Engineering Research 2020.
- [13] Singh, Satyendra & Sharma, Nitish & Chandrakant, Shukla & Singh, Surendra. " Electric Vehicles in India",2021.

- [14] A. Gaurav and A. Gaur, "Modelling of Hybrid Electric Vehicle Charger and Study the Simulation Results," International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET), 2020.
- [15] T. Payarou and P. Pillay, "A Novel Multipurpose V2G & G2V Power Electronics Interface for Electric Vehicles," IEEE Energy Conversion Congress and Exposition (ECCE), 2020.
- [16] A. V. Makarov, T. V. Makarova, V. G. Makarov and D. O. Kopp, "Optimal Driving Control of Hybrid Electric Vehicle at a Given Torque," 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), 2020.

Categorization of Fruit images using Artificial Bee Colony Algorithm based on GLCM features

Akshay S

Department of Computer Science
Amrita School of Arts and Sciences
Amrita Vishwa Vidyapeetham,
Mysuru, Karnataka, India
s_akshay@my.amrita.edu

Deepika Shetty M A

Department of Computer Science
Amrita School of Arts and Sciences
Amrita Vishwa Vidyapeetham,
Mysuru, Karnataka, India
deepikashetty1973@gmail.com

Abstract— The manual recognition of observed items from the records, which takes a long time. We suggest a neural network approach to classify apple defects. Due to the various fruit varieties classification is a challenging task. We suggested a system of classification based on an artificial bee colony algorithm (FSCABC) and a neural feeding network to classify fruit more precisely (FNN). First, fruit photographs were purchased using a digital camera and then, using a break and fuse algorithm, the backdrop of each image was stripped. To catch the fruits, we used a square window, where we downloaded the images in 256 x 256. Second, each fruit image was derived from a color histogram, texture and shape features to make a feature space. Thirdly, a major components method was used to minimize the dimensions of function space. The function is reduced and the weights and preferences are eventually trained on the FNN of the FSCABC algorithm. We have used the K-fold cross-validation approach to improve the capacity of the FNN. Test findings revealed that the SAFCABC-FNN achieved a precision of 89.1 percentage in 1653 fruit imagery in all 18 categories. The rating accuracy was more than the GM-FNN (GA-FNN) with a rating of 84.8 percentages, 87.9 percentages, the PSO-FNN with 85.4 percentages, the ABC-FNN with 88.2 percentages and the supporting kernel vector. Then FSCABC-FNN was regarded as effective in the fruit classification.

Keywords—Feed Forward Neural Network, Fruit classification, Channel Separation, K-Means

I. INTRODUCTION

Classification of fruit is an important and daunting process for supermarkets, since in order to establish their price the cashier has to know the fruit types of a given fruit. This dilemma has been largely overcome with the use of barcodes on packaging goods but most people continue to select their products themselves. These fruits cannot be stacked with the bar codes and must thus be weighed. One approach is to issue fruit code; however, tired barcode storage will lead to errors in pricing. Another option will be to store the cashier with picture and password, but the inventory booklet will take time to flip (Rocha et al., 2010). In the past decade, scientists have proposed several promising solutions to the same issue based on computer vision and machine learning. First fusion of data by Baltazar and al. (2008) along with an instructively fresh and untreated tomato in a three-class Bayesian classification.

Primary component analysis (PCA) for the bruised long fruits, partial lower-square discrimination methods and soft-independent class analogy modelling to construct classified models. Veggie Vision was a supermarket integral size, image system and user friendly interface. The morphological analysis was used for Hong et al. (2006) to distinguish walnut and hazelnut into three classes. selected Max Wins to vote for their Gaussian kernel RBF in order to examine the different forms of fruits. For a computerized system has been set up to assess the consistency of impurities in sample oils. The two

hidden layers of a backbone network were employed by Fan et al. (2013) for modelling the structural characteristics of extrusion photos of food on the superficial food surface. Intelligent framework designed by Omid et al. (2013), using hybrid fluffy logic and computer viewing methods for egg grading using criteria as egg defects and dimensions, was developed. One or more of the following 4 shortcomings can arise from the techniques described above. You first require external machines such as a gas machine sensor, an light machine sensor or a weight machine sensor. Second, it is not sufficient for all fruit to have the classification systems, i.e. The third author, as many pictures of a fruit may have similar or identical colour and form properties, may not have rigorous recognition systems. Fourthly, customers say that the results are incorrectly graded. This research paper seeks to propose a modern computerized fruit image classification project with the goal of solving four main deficiencies. First of all, we all use only a digital camera to clear other machine sensors that are complex. Second, as many varieties of fruit as possible should be recognized in the suggested classifier.

This research includes 18 kinds of fruit. Thirdly, we not only capture traditional form and colour, but we also capture prominent texture characteristics. Finally, the suggested classification will be accurate by the FNN and it would be an effective tool for supervising classifiers because it can randomly classify non-linear divided style and estimate an ongoing feature. In addition to the training for FNN various global optimization algorithms have recently become available. The BP, GA, SA and PSO algorithms all require unfortunately high calculation costs and can easily be stuck into the best of the local ones. This suggests that the outcome is likely to end without the best weights / preferences of the FNN. In this analysis, we use the algorithm for the optimum FNN weight/biases of the artificial Bee Colony (ABC). Karaboga et al. (2004) originally proposed ABC algorithm based on a group action of honey bees that are better performative than the GA, differential evolution and PSO in function optimization problems (Karaboga and Basturk, 2008). We realize that only one search procedure per iteration is conducted in traditional global optimization techniques. The PSO conducts, for instance, an initial global quest and a final local search (Zhang and Wu 2011), in addition, the ABC is in the position to undertake worldwide and national search in all sections. The probability to find maximum is significantly increased and local optima is prevented ultimately. To maximize the effectiveness of ABC and call the new method a chaotic fitness artificial bee (FSCABC), This paper author has used exercise and chaotic theory as the basis of the ABC (Zhang et al., 2011b). The following article has following structure: The findings are processed in Section 2, including pre-processing, extraction and reduction of features and stratified cross-validation processes. Section 3 defines the FNN classifier, introduces FSCABC methods of Algorithm to

train the biases of F-Neural Network, and summarises the steps taken by system for observation of fruits. Section 4 represents the experimental outputs one by one. The FSCABC/FNN outputs apply to GA-F-Neural Network, PSO-FNN, ABC—F Neural Network, and a method for supporting the kernel variable (kSVM). The conclusions were discussed in Paragraph 5. Final Section 6 deals with the conclusion and future research.

II. LITERATURE REVIEW

Maximal Entropy Random Walk for Region-Based Visual Saliency in SEPTEMBER 2018 by Jin-Gang Yu, Student Member, IEEE, Ji Zhao, Jinwen Tian, and Yihua Tan, Member, IEEE. [1] They proposed a novel bottom-up model for detecting salient materials in natural pictures. They were encouraged by the now-a-days advance in the statistical thermodynamics, they acquired a novel problem based architecture, the maximal entropy random walk to calculate saliency. In sum of, as a novel saliency measure, one would also be interested in MERW's performance for the work of eye-fixation prediction. Badly, while the overwhelming majority of existing architects are space bases, region bases saliency has not at all correctly addressed.[1] Random Walk and Graph Cut for Co-Segmentation of Lung Tumor on PET-CT Images in DECEMBER 2017 by Wei Ju, Dehui Xiang, Bin Zhang, Lirong Wang, Ivica Kopriva, Senior Member, IEEE, and Xinjian Chen*, Senior Member, IEEE. [1] They effectively merge 2 methods by doing fully use of the superior contrast of PET pictures and superior spatial resolution of CT pictures. Random walk and graph cut method is merging to resolve the segmentation issues, in which random walk is used as an initiate tools to give materials seeds for graph edit segmentation on the PET and CT pictures. To smaller the energy methods on the constructed graph to give an optimal segmentation. Although PET-CT pictures have been broadly utilized in lab, automatic segmentation on PET-CT pictures are still very difficult task.[2] Guided filtering for p r n u-based localization of small-size picture in 2018 by Giovanni Chierchia, Davide Cozzolino, Giovanni Poggi, Carlo Sansone, Luisa Verdoliva. The present or absent of the camera PRNU type is calculated by a correlation validation. Given the very low energy of the PRNU signal, however, the correlation must be averaged over a pretty large window, reducing the algorithm's power to expose tiny forgeries. To increase resolution, they prepare for correlation with a spatially adaptive filtering method, with weighing computed over a suitable pilot picture. Implementing efficient is achieved by resorting to the newly implemented guiding filters. Process difficulty is higher.[3] Link predicting the power of maximum entropy random walk in 2014 by Rong-hua li Chinese university Of Hong Kong. The ancient random walk (TRW) includes the link architectures by treating all nodes in a network equally, It is ignoring the central nodes of network. In many real-time network, nodes of the network not only preferred to connect to the same type of node, but also prefer to connect to the middle nodes of network. It performs all the other unsupervised methods as well as the supervised approaches. The power of the procedure is future to be improvement.[4] Splicing forgeries localization through the utilization of 1st number features in 2019. A approach to differentiate & then localize a single and a double image compression in divisions of a picture from the utilization of the DCT co-efficient 1st number components and supporting a support vector machine (SVM) classifier is implemented. The proposed method will

be also compared with a different procedure approach, based on the analysis on the image artifacts invented by the double image compression where an automatic algorithm to localize the forgery is architecture. Regarding duplicates individuation 3 are the principle objects of detectors observed so far.[5] A Bayesian-MRF method for PRNU-based picture duplicate forgery identification in 2019 by Giovanni Chierchia, Giovanni Poggi, Carlo Sansone. They introduced a method to identify picture duplications utilizing sensor type noise. Casting the problem in terms of Bayesian methods. They are utilizing an Markov random field prior to architecture the strong spatial dependencies of the source, and take decision effectively on the full picture other than individual for every pixel. It was were improved upon the seminal PRNU-based duplication identification method introduced but less accuracy. Localization of maximal entropy random walk in by [6] they are defining a new class of random walk processes which makes maximize entropy. This method maximal entropy random walk is equivalent to generic random walk if it takes place on a regular lattice, but it is not if the underlying lattice is informal. This localization phenomenon, which is purely classical in real, is detailed in terms of the states of a certain random function. The lattice is irregular.[7]

III. EXISTING METHODOLOGY

A. Architecture of neuron

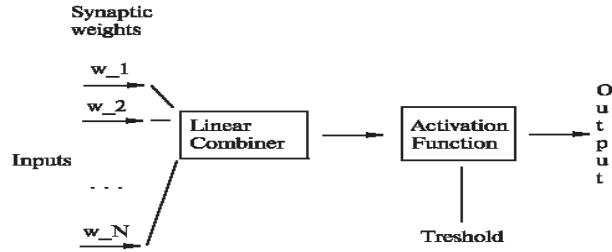


Fig.1. Structure of a neuron

"Artificial" neuron would have one input (all Neuron). The neuron has nodes, as you see, which are called synapses: these nodes also have nodes that connect them to the inputs, outputs, or some other neuron. A single linear unit that uses a value-taking algorithm for all the inputs. A simple solution is to apply dInput together (if you are not a programmer, where "d" means "double," we use it as a synaptic dWeight-multiplier of the floating-point number (dinput):

```

for(int i = 0; i < nNumOfInput; i++)
    DSum = dSum + DInput[i] * DWeight[i];
  
```

Take this situation into account - the human ear could work near the jet engine, and at the same time - if it were 10 times more susceptible, we could feel a single molecule entering our ears on the membrane. What's the significance of this? That does not necessitate linear feedback. The disparity between 100 and 200 could be equal if 0.01 or 0.02 is to be obtained. For a nonlinear way mechanism of activation is used. It takes NEVER input from less endlessness to more endlessness, and pushes into -1 to 1 or 0 to 1. A neuron not just a highly effective mechanism for pattern recognition. If we integrate neurons in the multilayered system known as

neural networks it takes into consideration the real strength of neural networks. We have 3 layers in our network (we can do more, but we would have a lower network if we do less). You can also make 4 levels if you look for unclear objects. And I never saw a crisis with five layers. For 99 percent of functions, 3 layers are the best option. In the first layer, there are N -neurons where N is equal to the input number. The output layer comprises M neurons with M being identical to the output number. Multi-layer feed-forward networks. There's a layout in a feedback network. The units in each layer collect their entry from units immediately below the layer and transmit their outcome to units directly above the unit in the layer. Inside a layer there are no ties. The inputs of N_i are fed to $N_{h,1}$'s first plate. The input drives are just 'fan-out' units; they are not processed. The "Fi" operation of the weighted inputs and the triggering of a hidden units is equally significant.

$$Y_k(t+1) = F_k(S_k(t)) = F_k(\sum_j W_{jk}(t)y_j(t) + \delta_k(t)) \quad (1)$$

The cached output would be allocated to the next N_h line; 2 to the final layer of cached output units to N_o output unit layer.

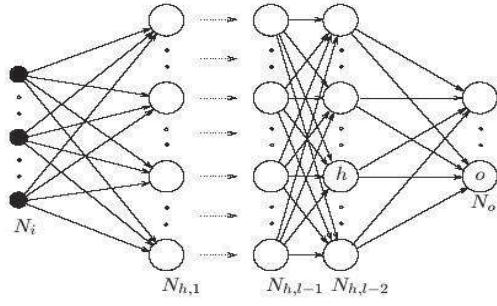


Fig.3. Multi-layer feed-forward networks

Back propagation to all layers of networks can be expanded, but only a layer of hidded units can accurately determine operation with a finite number of connections, provided the activation function is necessary. (the uniform solution theorem). A buoyant network, with only one concealed interface layer, has a sigmoid triggering feature for most programme implementations.

B. Understanding backpropagation

The equations may be mathematically right in the last chapter, but what do they mean? May we explain history's distribution instead of rectifying the equations necessary? Naturally, the answer yes. In reality, whole mechanism of background diffusion is very obvious. In the above-mentioned equations what happens is the following. The results are translated to the outlet units while the testing trend is screened, with the real network output being compared to the predicted output values. For a certain output unit let's mark the error e_o . e_o must be brought to 0. The simplest solution is by the greedy method: the relationships in the neural network can be modified for this particular model

about zero every time. In Delta law, we realize that incoming weights must be modified to reduce errors.

$$\Delta W_{ho} = (D_o - Y_o) Y_h \quad (2)$$

That's the first move. But this isn't sufficient on its own: as we apply this rule alone, the scale of the forward network will never be changed and we have no absolute power of representation, as the universal approximation theorem promises it. Again, we want to apply delta legislation to change the weights from input to concealed units. However, which distributes the error of a production unit, weighed by this relationship, to all related hidden units. In contrast, hidden unit received a delta from output point, which is proportional to the weight of the relation between them multiplied by the delta.

C. Working with back-propagation

Two measures are taken to obey the generalized Delta rule: In phase one x is shown and spread across the network is to decide the position of Y^p_o for each output point. The outcome is $\delta_p = o$, an error signal related to each output unit's expected value. The 2nd stage describes reversing the network by transmitting an error signal to any network device and correctly calculating weight differences.

D. Weight adjustment with sigma activation function

Weight of the interface is equalized in proportion to the error signal product μ , which transmits this signal to receive input and output for the k unit j along the relation.

$$\Delta_p W_{jk} = \sqrt{\partial_p^k Y^p_j} \quad (3)$$

If the instrument is a production unit, an error signal would be given.

$$\partial_p^o = (d^p_o - y^p_o) F'(S^p_o) \quad (4)$$

The function 'sigmoid' as defined takes the activation function F :

$$Y^p = F(S^p) = 1/(1 + e^{-SP}) \quad (5)$$

To allow the error signal to be written for an output unit as:

$$\partial_p^o = (d^p_o - y^p_o) y^p_o (1 - y^p_o) \quad (6)$$

A device to which it is connected and the weight of its connecting is the error signal of the hidden unit repetitively calculated. For sigmoid activation mechanism. Learning rate and momentum A proportional weight adjust ion is required for the study protocol $\partial E^p / \partial W$ The real downward gradient implies a very small amount of activity. The ratios are the pace of learning continuously. For practical reasons, we choose the best possible learning rate without oscillation. One way of preventing oscillation in general is by using a term of

momentum to allow the difference of weight dependent on prior weight changes:

$$\Delta W_{jk}(t+1) = \sqrt{\partial p_k} y_{pj} + \partial \Delta w_{jk}(t) \quad (7)$$

In cases where t indicates the number and F is the effects of weight shift from the previous weight are constantly determined. While theoretically, the background propagation algorithm only leads to a downward gradient tendency towards a total error when weights have been adjusted according to the whole range of study models, often the learning rule applies separately to every design, i.e. uses a pattern p , calculates E_p and adjusts weights ($p=1,2,\dots P$). This leads to faster convergence, and an analytical indication also exists. However, we have to take note of the order in which the habits are learnt. For example, by using the same string, the network will focus on the first few patterns again and again. To overcome this problem, a permuted teaching approach is used. The main objective is to identify and segment the deficiencies in the given picture. The Local Binary Pattern algorithm, Colour histogram, removes features from the invariant colour picture. The defective apple can be recognized by the neural network. For image converter $L^*a^*b^*$ Colour Space. These images are divided by the mechanism of clustering. The fruits infected are categorized in Bayesian classification. Local binary pattern algorithm is used for extraction of characteristics. Any unchecked study algorithms are then used to identify the photographs of the fruit. The photos are categorized without targeting in non-supervised learning approaches. The same characteristics are clustered. The cluster id is returned from the cluster to which the picture matches. Otsu segmenting algorithm is used for segmentation. The classification accuracy is poor as only the binary algorithm of the local pattern is used for classification. The unattended approaches to learning are not effective grouping approaches.

IV. PROPOSED METHOD

In the first case, Apple images are preprocessed such that redundant pixels are removed. The photos with the local binary model's algorithm delete the attributes. Functionalities are retrieved and stored as training image attributes for all images in the database. For the reference image as test image functionality, the features are omitted and preserved. The training image and the test image attributes are transferred into the classifier. The neural network separates a malfunctioning picture into a natural picture. If the image is strange, the images are clustered by k-means into segmentation algorithms. The photograph type of disease then is detected by a multi-SVM classifier. Finally, it is measured the precision of the designation. Precision in the classifier means that the accuracy of the solution proposed is comparable with previous algorithms. The functions derived from local binary pattern algorithms are more powerful and accurate to improve the accuracy of the classification.

A. Module description

1) Preprocessing

Unnecessary noises in the picture were removed during preprocessing. The unwanted pixels in these frameworks are an unwelcome noise.

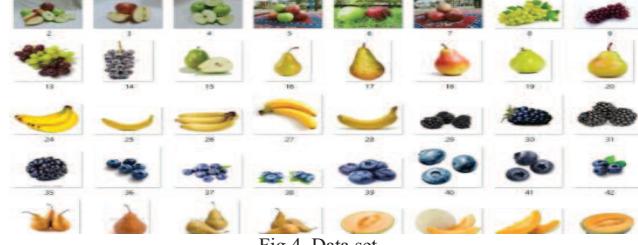


Fig.4. Data set

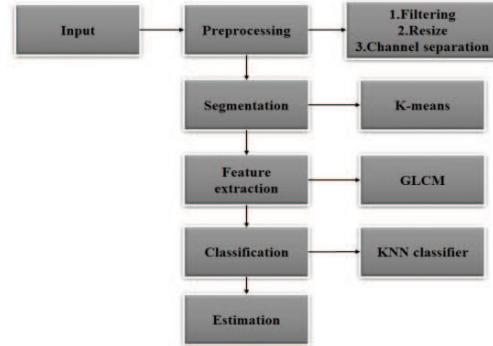


Fig.5. Architecture design

Preprocessing approaches of an input picture use a tiny pixel neighborhood in order to get a new brightness in output picture. These preprocessing processes are also referred to as filtration. Pre-processing local approaches should, in accordance with the processing objective, be split into two categories. The smoothing removes the noise or any minor image fluctuations equal to the removal of the high frequency domain. Smoothing unfortunately blurs any sharp point, too, with crucial picture detail. Gradient operators draw on local derivatives of image function. Derivatives are greater if the image's characteristics vary fast. The objective is to show certain locations within the gradient operator image. In preprocessing methods, filters are used. The multidimensional array noises removal feature is performed using `imfilter()`. Each class or dimension may be logical input or a non-sparse numerical array. The sequence that results has the same size and class as the input. Double precision floating point numbers are used to calculate. The noise of the images is filtered during the preprocessing. Noises in the image are needless pixels. For reduction of sounds, the Gaussian filter is used. This clarifies the image and all pixels in the clip. For using the Gaussian filter, we use `imfilter()`.

2) Feature extraction

The image removes the Color Coherence vector and local binary characteristics in order to extract characteristics from the Color Histogram image. On each color channel, the image color channel is separated by histogram. Functionality principles are held. The color cohesiveness vector is dependent on the image and the values are stored as functions. The difference between the pixels and the next pixels is then processed, as features gradually have LBP functions. A color histogram shows the color distribution of an image. A color histogram for digital pictures represents the number of color pixels that cover the color space of the image, and the collection of colors in each of the fixed color ranges list. The color histogram can be found in all color spaces, but for three-dimensional spaces such as RGB or HSV the term is used more frequently. The word histogram of intensity should instead be used for monochromatic

images. The color histogram is N-dimensional for multi-spectral images where each pixel is measured by the number of measures (for example, above the three measures in RGB). The number of measurements taken by N is N. Each measurement has its own light spectrum wavelength set, some of which can lie beyond the visible spectrum. The histogram is essentially the pixel count. When the possible color values are enough, each color is placed on the scale. Space is usually divided into a series of sets, often organized as a regular grid, each having a number of identical color values. A smooth feature, specified over the space approximating pixel counts, can also be displayed in the color histogram.

3) Initial classification

The pictures are categorized as normal or abnormal by the classification of the neural network. The NN gets inputs that can be any sort of pattern. If the neuron has obtained its input in the first layer, the linear combiner and activation function are applied to the inputs and the output is generated.

4) Segmentation

Clustering K-means is a method of vector quantification that has a strong influence on cluster analysis in image processing. Clustering K-means is designed to divide n observations into k classes, with each observation belonging to the nearest medium cluster, as a cluster model. Therefore, the pure k-mean algorithm is not quite scalable and restricted as a result. Many other algorithms have since been created for these applications. The regions are then clustered in a series of groups by means of k-means clusters. In the end, a pixel wise segmentation is added to pixels not segmented in the first step. With this double-step procedure, the computational cost can be greatly reduced, as the second stage requires only a small number of pixels to be segmented, preventing the estimation of the function for any pixels in the image. In addition, the first stage of the segmentation process is improved with parameters calculated from the regions in this picture. The region that is infected is separated by the abnormal images. The K-means algorithm is used for segmenting images. The picture of the fruits is modified by RGB to RGB in L*a*b*. The defecated regions are joined together into one division. Depending on the pixel values variations, the image is divided into classes.

5) Classification

The area of the Network is probably a boundary between artificial data and approximation algorithms. Consider this as an algorithm for the smart solution. The NNs are usually used in global approval, tools for getting from its environment, tools for identifying unapparent dependency on results, etc. The brain is a multilayer network about 10^{11} neurons (which is around a total of 6-7 neuronal layers, if it's the human cortex) that acts as a paralegal machine, able to learn from the world's "feedback" and adjust its architecture either by the development of new neural relations or ecological activities, which during the work could alter computer hardware. We will need to say that a normal neuron is always connected to 50-20100 of the other neurons in order to understand this figure. The brain consists of neurons that communicate, simply to put it. Classification is done using SVM classifier. Machine learning helps vector machines that analyses data and identify patterns used for analysis of classification. The basic SVM gathers input data and predicts which performance of the two possible classes renders a

binary classification unlikely. Given a number of training scenarios, each of which is marked as one of the two categories, an SVM training algorithm produces a new model for each category. A SVM model is a space-point example that is mapped to separate the examples of each type as simple and as far as possible.

V. EXPERIMENTAL RESULTS

A. Input Image

This is an input image; we will input an image to start all the processing to estimate what type of fruit it is.



Fig.6. Input Image

B. Split and Merge

After inputting an image first, we split it as black and white also called as binary image. Then we merge the image here we can see only the fruit image and all the unwanted parts are erased.



Fig.7. Split and merge

C. Image Resize

Image resize means here we convert unsized image to the standard size that is 214 X 214 pixels for the image processing easy.



Fig.8. Resized image

D. Filtered image

Next, we start removing the noise of the image using noise filtering here we use Gaussian filter to remove the noise of the image.



Fig.9. Filtered image

E. Channel Separation(Red, Green and Blue Band Image)

Next comes to channel separation which is done using kernel neural network algorithm, here we separate an image as red band images, green band images and blue band images.



Fig: 10. Channel Separation

F. Feature Extraction (Color, texture and shape)

Here we extract the features of the fruits that is its color, texture and shape using Principle component analysis.

	1	2	3	4
1	0.0189	8.9767e-04	0.0018	0.0001
	1	2	3	4
1	266.4098	308.7412	55743	313

Fig.11. Feature Extraction

G. Feature Reduction

Feature reduction is the transformation of information from high dimensional space into low dimensional space so that it is used for all the processing easy and understandable.

H. Classification

After all the processing we finally classify the type of the fruit using neural network here we use kernel neural network algorithm

Identified Fruit type "Apple" by NN classification

Fig.12. Classification output

I. Neural Network Results

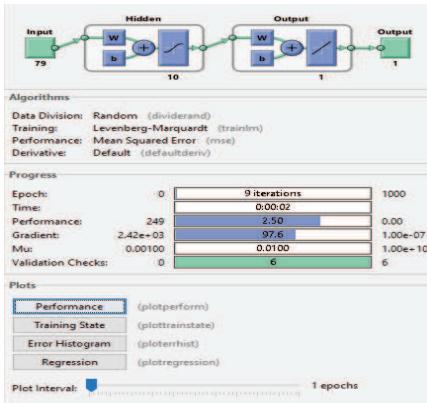


Fig.15. Neural Network

TABLE 1 MAXIMUM F – MEASURE TABLE OBTAINED FROM THE METHOD.

Testing/Training Ratio	50:50	60:40	70:30
Max F - measure	95.6	96.8	97.4

VI. CONCLUSION

We have an algorithm that segments the apple picture to specify the area which has been vanquished and to classify the image as extracted attributes. The sound in the pictures is eliminated in the preprocessing phase. The picture color channels are broken down and color history charts are added to each color channel. The derived characteristics are histograms calculated. Invariant color features are first derived and local algorithms are used. The neural network then discovers that the defect is present or not in the apple frame. The apple's photos are converted into L*a*b* color. The SVM classifier has defined apple defects, such as apple scab, apple red, or apple block, using the extracted features and the real name. The vector quantized data can be represented in a better way as a tree to make the matching faster. A similar kind of an attempt can also be made on

interval valued data representation. Indexing and hashing can be used to improve the results. These methods may be implemented for speech related research.

REFERENCES

- [1] J. Hartman, Harvest Fruit Disease, Colleges, University of Kentucky, http://www.ca.uky.edu/agcollege/plantpathology/ext_files/PP_PP_FShthml/PPFS-FR-T-2.pdf, as viewed in December 2011. The pathology sheets of this article were published in German.
- [2] The Agriculture Computers and Electrotechnology Computers, vol. 36, pp. 215–223, Nov. 2002: Q. Li, M. Wang, and W. Gu. "Computer vision based apple surface defect identification method."
- [3] MM. Mehl, K. Chao, M. Kim, and Y.R. Chen, 'Applied engineering of agriculture, vol. 18, pp. 219–226, 2002. Applied agricultural engineering, vol. 218–226.
- [4] M. S. Kim, A. M. Lefcourt, Y. R. Chen, Y. Tao, "Multi Spectral Image Fusion Automated identification of Apple Fiction," Food Engineering Journal, vol. 91, 2005, pp. 85-91.
- [5] O. Leemans, & M. F. Destain, O. Leynen, & V. Leemans. Development of multi-spectral perception sensing devices for apples. The diary of food technology, vol. 69, pp. 41-49, 2005.
- [6] OFF. V. Leemans, H. Magein and M. Fr. Destain. Agriculture Computers and Electronics Vol. 23, 43 to 53, June 1999. The Bayesian vision and classificatory defeat segmentation scheme of jonagold apples.
- [7] V. Leemans, H. Magein, and M. F. Destain, Computers & E-Systems in Agriculture, vol. 20, pp. 117–130 of July 1998, "Defect Segmentation of 'Golden Delightful' Apples Using Colour Machine Vision."
- [8] Pattern Analysis and Machine Intelligence, "Local Binary Pattern Recognition for Gray-scale Multiresolution and Invariable Rotation Texture"
- [9] "Outex – a new mathematical analysis paradigm and texture analytical algorithm,". Pattern Recognizing Conferences 2002, p. Mäenpää, J. Viertola, J. Kyllönen and S. Huavinen. Ojala and J. Viertola.
- [10] Ahonen, T. Hadid and Pietikäinen, IEEE Trans. on Machine and Master Analysis, Vol. 28, no.12, pp. 2037-2041,2006: "Eye reconnaissance with local binary patterns.
- [11] "The dynamic textural sensoring uses bilateral local designs with the implementation of facial expressions," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 27, vol. 6, pp. 925-938, 2007;
- [12] X. Huang 10. X. Z. Li, S. "Location of shapes dependent on mathematical approaches by generalised local binary designs," X. Huang, Y. Wang, in ProC. International Conference on Image & Graphics, 2004, pp.184-187.
- [13] Z. Guo, L. Zhang, and D. Zhang, IEEE, Trans. on Image Processings vol. 19, no 6, pp. "The Local Binary Operator Complete Texture Classification Model.
- [14] 'Some ways to characterise and analyse multivariate results' L. M. LeCam & J. Neyman: vol. 1, pp. 281–297, Fifth Berkeley Symposium for Mathematical Statistics and Probability, vol. 1.
- [15] R. Gonzalez, R. Woods, 3rd ed., Prentice-Hall, 2007. Image Processing In ACM Multimedia, 1997, G. Pass, R. Zabih and J. Miller, "Compare image use vectors for colour coherence." Page 1 to 14.
- [16] A. Rocha, C. Wainer, and D. Siome, A. Rocha, C. Hauagge, Agriculture Computers, Elsevier; Vol. 90, pp. 96-104, 2010, "Automatic picture recognition for fruit and forests
- [17] Akshitha Raj, R., Gopika, M., Apoorva, P., Akshay, S., "Segmentation and classification of fruit images independent of image orientation using height width vectors", 2019, International Journal of Innovative Technology and Exploring Engineering, 8 (9), pp. 3232-3237.
- [18] Akshay, S., Mandara, S., Rao, A.G. "Facial expression recognition using compressed images", 2019, International Journal of Recent Technology and Engineering, 8 (2), pp. 1741-1745.
- [19] Amrutha, C. V., C. Jyotsna, and J. Amudha. "Deep learning approach for suspicious activity detection from surveillance video." In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pp. 335-339. IEEE, 2020.
- [20] Krishnan, Sajitha, J. Amudha, and Sushma Tejwani. "Gaze Fusion-Deep Neural Network Model for Glaucoma Detection." In Symposium on Machine Learning and Metaheuristics Algorithms, and Applications, pp. 42-53. Springer, Singapore, 2020

Detection of Falsified Selfish Node with Optimized Trust Computation Model In Chimp –AODV Based WSN

¹P.Manjula

Department of Information Technology

Veltech Multitech Dr.Rangarajan Dr.Sakunthala Engineering College,

Avadi, Tamil Nadu.

manjula.arunraj@gmail.com

²Dr.S.Baghavathi Priya

Department of Computer Science and Engineering,

Rajalakshmi Engineering College,

Rajalakshmi Nagar, Thandalam, Tamil Nadu.

baghavathipriya.s@rajalakshmi.edu.in

Corresponding Author: P.MANJULA

ABSTRACT

In Wireless Sensor Networks (WSNs), energy and security are two critical concerns that must be addressed. Because of the scarcity of energy, several security measures are restricted. For secure data routing in WSN, it becomes vital to identify insider packet drop attacks. The trust mechanism is an effective strategy for detecting this assault. Each node in this system validates the trustworthiness of its neighbors before transmitting packets, ensuring that only trust-worthy nodes get packets. With such a trust-aware scheme, however, there is a risk of false alarm. This work develops an adaptive trust computation model (TCM) which is implemented in our already proposed Chimp Optimization Algorithm-based Energy-Aware Secure Routing Protocol (COA-EASRP) for WSN. The proposed technique computes the optimal path using the hybrid combination of COA-EASRP and AODV as well as TCM is used to indicate false alarms in detecting selfish nodes. Our Proposed approach provides the series of Simulation outputs carried out based on various parameters

Keywords: TCM, COA, WSN, COA-EASRP, AODV

1. INTRODUCTION

A Wireless Sensor Network is a collection of sensor nodes that can be used for either broadcasting or multicasting communication. WSN consists of a group of sensor nodes connected wirelessly which is used to send and receive the data of the environment. It Organizes the sensed data at the sink node. WSN is used to measure the environmental conditions like sound, wind, humidity, temperature,

pressure, etc., the network is bi-directional where sending and receiving of data can be done simultaneously. It also controls the sensor activities. WSN is an infrastructure-less network in which sensor nodes are considered not to be secured with limited available energy, whereas the base station is considered to be secure with unlimited available energy. The unattended working environment of WSN contains several weak points that affect adversaries.

Wireless Sensor network also comes with constraints. Since wireless, the network is not secured. There are numerous ways for an attacker to attack a WSN. An intruder can be able to overhear the various fields of a message in transit and alter the original message of a sender. The recipient receives the intruder's message not the original message sent by the sender. The transmission of data is based on the lifetime of the battery. The communication path has been changed if any node dies in the network. So, recharging of the node is necessary at regular intervals of times. Compared to a wired network, a wireless network has a low communication speed. Clustering is a technique used to optimize the energy exploitation in the nodes of the WSN. Cluster-dependent communication is now favored. Cluster-based communication, however, introduces extra overhead and pressure to the Cluster Head (CH) in dense network situations, which inevitably contributes to delays and hinders network performance. Energy conservation and security play a crucial part in transmitting data between the cluster head and sink node. The resource limitations of WSN sensor nodes make it difficult to communicate between sensor nodes, between a base station and sensor nodes, and between all sensor nodes.

Some of the characteristics of WSN are

- Resilience
- Node mobility
- Power consumption
- Scalability
- Ease of use

To defend WSN against malicious nodes, a variety of cryptographic and authentication-based routing approaches are available. Most modern cryptographic algorithms, on the other hand, necessitate a significant amount of computational overhead, power consumption, memory requirements, and communication bandwidth, making them unsuitable for resource-constrained (low memory, energy) WSNs. Network congestion is caused by the exchange of keys between nodes. WSN security will be rendered useless once keys in cryptographic techniques are leaked. To work well, most current encryption and authentication systems need centralized administration, which is typically not possible with WSNs. As a result, these encryption and authentication technologies are both cheap and ineffective. Security issues must be taken into account since sensor nodes are frequently powered by batteries. Energy consumption will increase if the authentication method is enhanced. Alternatives based on trust have evolved.

The main contributions of this paper are given in three steps.

1. A hybrid combination of COA-EASRP & AODV protocol is used for optimal path selection.
2. The Optimized Trust computation model is designed to indicate false alarms about the selfish nodes.
3. The performance of the proposed approach is compared with existing approaches in terms of Average End-to-end Delay (EED), Average Throughput and maliciously dropped packets in the routing layer.

The rest of this paper is arranged as follows. Section 2 gives the survey of Literature works. The problem statement and the methodology along with the algorithm are detailed in Section 3. Section 4 described the workflow of the entire paper. Section 5 details the simulation results. Section 6 concludes this paper.

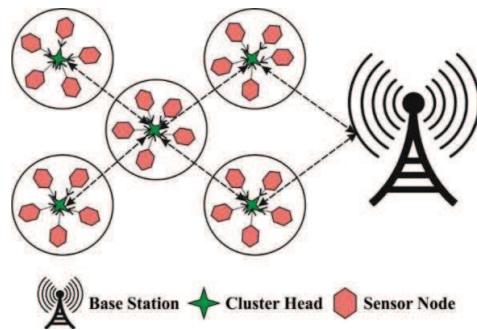


Fig. 1-WSN Structure

2. RELATED WORKS

In the present remote sensor organizations, security and energy are the most basic issues. WSN is defenseless 100% of the time to assault because of its open nature and unregulated arrangement, just as sensor node imitations. Insider packet drops are one of the most perilous assaults against WSN. A few cryptography-based innovations have been created to recognize and stay away from such assaults. WSN security will be rendered ineffective if keys in cryptographic approaches fail. The Watchdog is a conspicuous checking component in this procedure. Khishe, M.R. Mosavi,[1] suggests Chimp Optimization Algorithm (COA) inspired by chimps' individual intellect and sexual drive in collective hunting, which distinguishes them from other social predators.

Wang, Y., Zhang, M. & Shu, W [2] presented a reliable trust perception assessment model which can estimate a node's trust value based on its behavior and efficiently identify or isolate malicious nodes. The penalty and regulator functions are used to represent the impact of state changes on the trust value based on the node's conduct during the communication process. T. Khan *et al.* suggested method (LTS) operates on two levels, intra-cluster and inter-cluster, using a distributed and centralized approach to provide proper decisions about sensor nodes. Authors Tayyab Khan and Karan Singh suggested a trust-based reliable routing strategy to defend against selfish nodes using a hybrid trust model. This model makes use of a multifactor routing methodology with trust scores of nodes, residual energy, and path length as parameters. Selfish nodes also refuse to forward or postpone RREQ data and RREP conversations. The performance of WSN routing is significantly impacted by selfish node behavior. The divergence from the initial forwarding and routing is referred to as node misbehavior. As a result, a selfish

node identification strategy is necessary for MANET to encourage node cooperation. The most well-known and extensively used approach for identifying selfish nodes is Watchdog [11]. The relationship between AODV and a local watchdog can help with route recovery [12]. False positives and false negatives, on the other hand, are claimed to induce the watchdogs to create false positives and false negatives. Atakli et al. [13] present a weighted trust evaluation approach.

4. PROPOSED METHODOLOGY

The Proposed model has developed a COA-EASRP technique has been developed to derive optimal routes in WSN and thereby accomplish security. Besides, the COA-EASRP technique has derived an objective functioning involving various input variables such as namely energy, trust, and delay. The route with maximum energy and trust with minimal delay will be chosen as an optimal route by the COA-EASRP technique.

A. Process involved in COA

The COA-EASRP technique has derived a fitness function involving various input parameters to select routes in the WSN. The study considered a routing issue in WSN inspired by event-driven application. Based on the issue, event occurs randomly all over the sensing region. The sensor in the influence region of event, $R(H_i)$, identify the event and choose a leader (called as source node) amongst others, that could transmit the information to sink node. The source node should define a route in the network which

A1gorithm 1: Pseudo-code of COA algorithm

```

Inputs: The population size N and maximal amount of iterations t
Initialization of the arbitrary population  $X_i(i = 1, 2, \dots, N)$ 
while  $t < \text{max.amount of iterations}$  do
    for all the chimps do
        Determine the chimp's group
        With utilizing its group approach for updating
            end for
        for all search climbs do
            if  $x < 1$  then
                Upgrade the place of present search chimp
            else if  $x > 1$  then
                Choose an arbitrary search chimp
            end if
        Upgrade the place of existing search chimp
        end for
        Upgrade X Attacker, Barrier,
        Driver, and Chaser  $t + 1$ 
    end while

```

results in balanced energy utilization amongst the

increased QoS, sensor nodes, and improved privacy of data while some nodes from the networks have cooperated.

B. Application of COA for Optimal Route Selection

The COA-EASRP technique has derived a fitness function involving various input parameters to select routes in the WSN. The study considered a routing issue in WSN inspired by event-driven application. Based on the issue, event occurs randomly all over the sensing region. The sensor in the influence region of event, $R(H_i)$, identify the event and choose a leader (called as source node) amongst others, that could transmit the information to sink node. The source node should define a route in the network which results in balanced energy utilization amongst the increased QoS, sensor nodes, and improved privacy of data while some nodes from the networks are cooperated.

Assumed a multi-hop WSN characterized as (V, E) , finding of event trigger the sender node to define a routing path RP:

$$RP = (v_1, v_2, \dots, v_n). \quad (1)$$

Where v_n indicates the sink node, v_1 represent the source node, and $v_{i+1} \in FN(v_i), \forall i, 1 < i < n$ FN is collection of neighbors as follows

$$FN(v_i) = \{v_j | v_j \in Nbr(v) \&& dist(v_j, Sink) \leq dist(v_i, Sink)\}. \quad (2)$$

The following three objectives of the RP need to satisfy.

The cost function projected by Ok et al. [16] is adapted to calculate TEC. The EC transmit data packets from node v_i to v_j as follows

$$EC_{ij} = \frac{\text{Energy needed from node } v_i \text{ to } v_j}{\text{Remaining energy at node } v_i} = \frac{e_{ij}}{RE_i} \quad (3)$$

The sensor's aim is to choose amongst itself and its neighbors the optimal candidate for directly communicating to the sink by assuming the overall energy needed to communicate the sink node via the neighboring node. The TEC_{ij} of a neighboring node v_j at sensor v_i expressed by

$$TEC_{ij} = EC_{ij} + EC_{jSink}. \quad (4)$$

By utilizing TEC , node v_i decide either to directly interact with the sink or to transmit the information via node v_j .

When $TEC_{ij} < EC_{iSink}$, v_i sends its information via v_j , or else, it directly transmits the information to the sink.

The delay on the RP depends on the delays incurred at a node and the amount of hops in the path. Select a node i.e., near the BS amongst the neighbors of a transmitter node resulting in small amount of hops on the path. But, the node nearer to the sink might result in long queues or death of the node leads to additional delays in transferring the information. The source node v_i estimates the delay at a neighboring node v_j via path and node delay factor ($PNDF_{ij}$)

$$PNDF_{ij} = \frac{dist(v_j, Sink)}{dist(v_i, v_j)} * d_j \quad (5)$$

In which d_j represents the delay incurred by a packet when waiting in a queue at node v_j and is associated with the size of presented buffer space at node v_j . In the beginning, d_j is considered zero.

Node capturing is an extensively employed technique for an inside attack in WSN, i.e., the most laborious technique of attack to be prevented, because of the legal identity of the compromised node. In the study, the packet generation rate and packet drop rate of nodes are taken into account to calculate their DTF. The node reliability can be estimated by a DTF method,

$$DTF = w_1 * \left(\frac{p_r + p_g - p_t}{p_r + p_g} \right) + w_2 * \left(\frac{p_g}{p_{\max} r} \right) \quad (6)$$

Whereas w_1 and w_2 represent the weights, that indicate the significance of packet transmitting and dropping behaviors, p_r denotes the number of packets received by a node, r signifies the number of rounds, p_t indicates the number of packets transferred by that node, p_{\max} represents the maximal amount of packets a node could transfer with the presented energy, p_g shows the number of packets produced, and that is almost equivalent to

$$p_{\max} = \frac{E_{initial}}{P_{size} * (1.5 * E_{elec} + \varepsilon_{fs} * d_0^2)} \quad (7)$$

The reliability of nodes is inversely proportionate to the DTF values, that is, the lesser the value of DTF, the greater is the reliability of nodes [17]. For finding the suitable forwarding node and assessing the neighbor of the sender, the 3 purposes (that is, minimizing PNDF, maximizing DTF, and minimizing TEC) are integrated to an individual heuristic called selectivity value of the node (SVN), by relating weights with the above three purposes. The source

node v_i , the $v_j(SVN)$ is estimated by the following equation

$$SVN_{ij} = w_e * TEC_{ij} + w_q * PNDF_{ij} + w_s * DTF_j \quad (8)$$

Here, w_q denotes the weight of QoS parameter, w_e represents the weight related to the energy parameter, and w_s shows the weight related to the security parameter thus $0 \leq w_e, w_q, w_s \leq 1$ and $w_e + w_q + w_s = 1$. This weight represents the significance of this parameter and is altered based on the requirement of the application.

4. Problem with Existing TBS Approach

A monitoring node X examines its neighbor node Y depending on how reliably Y forwards packets received from X in existing TBS. If node Y misses packets, X has no way of knowing whether this is due to a network fault or malicious behavior on Y's part. If a packet is lost or dropped due to the EMI or collision effect or fading effect in a network which is occurred in node Y, Node X declared Y as Selfish Node S_n. Node Y will not be included in the optimal route and data will be transmitted to another trustable node. This is referred to as the false alarm problem. Because of the false alarm problem, the monitoring node removes some non-malicious nodes from the network.

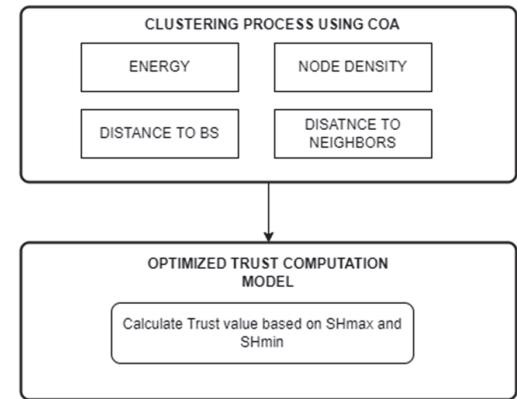


Fig. 2-Proposed Workflow

A. Optimized TCM Algorithm

Our key contribution to resolving the problem of false alarm is to provide wrongly discovered selfish nodes a second chance to be evaluated as good nodes. If a Tn is dropping packets owing to a network fault, it becomes an

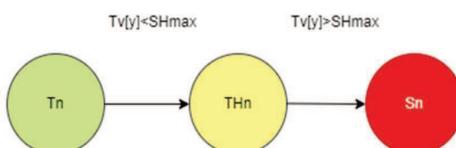
S_n in the present system. If a node has become S_n , then it cannot be changed back to T_n . As a result, we'll need a way to restore mistakenly recognized as good nodes in the routing process. In our proposed TCM model, a new intermediate state T_{Hn} is introduced between T_n and S_n to solve this problem. Figure 3 depicts the state transition of this strategy. In order to identify a node in one of the 3 States, TCM model will compute two thresholds, SH_{max} and SH_{min} (T_{Hn} , T_n , S_n) where SH_{max} is an upper trust threshold, and SH_{min} is a Lower trust threshold value.

Algorithm 2: TCM algorithm

Start

1. Source Node [SN] sends the Route Request Packet (RREQ) packet to all of its one-hop neighbors.
 2. Every typical neighbor node either rebroadcasts the RREQ to its neighbor node or sends a RRC packet to the Source Node in the event that it has effectively rebroadcasted a similar RREQ previously.
 3. After sitting idle for a prefixed timeframe, the source node checks its routing table and analyzes the conduct of its neighbors.
 4. IF the source node gets RREQ or RRC Packet in response from its neighbor node, THEN this neighbor node is characterized as normal node T_n .
 5. ELSE Source Node monitors node Y and calculates its trust value $Tv[Y]$
 - If $Tv[Y] < SH_{max}$ then
SN marks this node as T_{Hn}
 - Else If $Tv[Y] < SH_{min}$ then
SN marks this node as S_n
 6. END IF
 7. Flooding of RREQ Continues.
 8. Repeat the Steps 2 to 6 for each node in between the Source and Destination.
 9. Process will terminate after evaluating all the nodes.
- End

Fig. 3-State Representation of Optimized Trust Computation Model



Assume that node X wants to transmit an RREQ packet to node Y within t seconds. Node X determines Y's trust value $Tv[Y]$ based on firsthand observations of Y's packet forwarding behavior over t . $Tv[Y]$ is determined using the network's Beta trust model, which is worth

highlighting. Node Y will be classed as T_{Hn} rather than S_n if its trust value falls below SH_{max} (as in existing trust system). It's categorised as T_{Hn} since it's probable it's not a malicious node. It is losing packets due to a network problem, decreasing its trust value.

5. Simulation Results

At the routing layer, malicious packets that have been dropped are counted. Figure 4 shows the performance comparison in terms of maliciously dropped packets at the routing layer. TCM's proposed approach resulted in fewer packets being dropped maliciously.

Fig.4-Maliciously Dropped packets at the routing layer

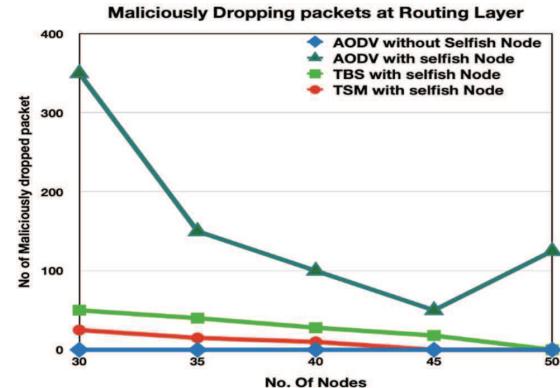


Figure 5 depicts the average throughput. When compared to earlier methodologies, our proposed methodology has a greater average throughput. The proposed technique has a throughput of 50 kbps at 30 nodes and 68 kbps at 50 nodes. Traditional AODV has a throughput of 70 kbps at 50 nodes, which is nearly equal to the suggested technique. Traditional AODV and TBS have throughputs of 43.5 kbps and 58 kbps at 50 nodes, respectively, which are lower than the suggested technique. Figure 6 depicts the suggested methodology's average end-to-end delay (EED). Traditional AODV with selfish nodes has a larger EED than the proposed technique.

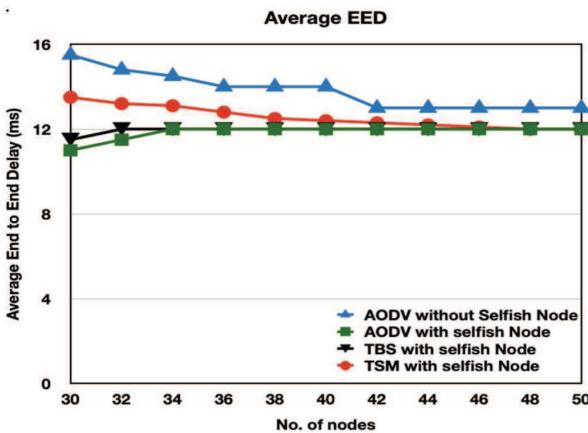
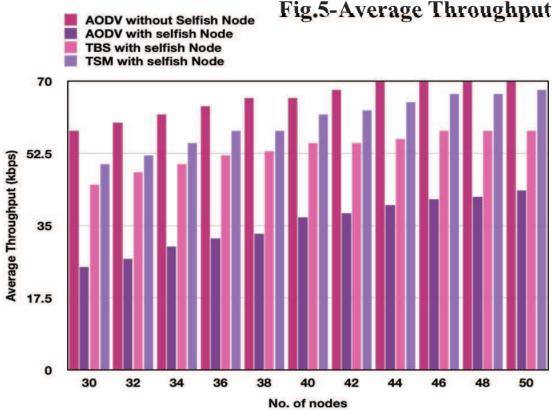


Fig.6-Average END TOEND Delay

6. Conclusion and Future Work

In this paper, our proposed approach developed a TCM model for detecting falsified selfish node alarms in order to improve the energy efficiency of WSN. Chimp optimization-based clustering provides efficient energy consumption in WSN. As a result, the proposed method implemented our TCM algorithm in a chimp optimization cluster-based environment along with AODV for optimal routing. In comparison to existing TBS and other approaches, our suggested technique improves throughput as well as detecting falsified selfish nodes in WSN. For the computation of trust values, the Beta trust model is used in this work. In the future, work will be enhanced with secure data aggregation in WSN.

7. References

- M. Khishe, M.R. Mosavi, Chimp optimization algorithm, Expert Systems with Applications, Volume 149, 2020, 113338, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2020.113338>
- Wang, Y., Zhang, M. & Shu, W. An emerging intelligent optimization algorithm based on trust sensing model for wireless sensor networks. *J Wireless Com Network* 2018, 145 (2018).
- T. Khan *et al.*, "A Novel and Comprehensive Trust Estimation Clustering Based Approach for Large Scale Wireless Sensor Networks," in *IEEE Access*, vol. 7, pp. 58221-58240, 2019..
- Tayyab Khan ,Karan Singh" TASRP: a trust aware secure routing protocol for wireless sensor networks" International Journal of Innovative Computing and Applications 2021 Vol. 12, No. 2-3.
- Dietrich, Isabel, and Falko Dressler. "On the lifetime of wireless sensor networks." *ACM Transactions on Sensor Networks (TOSN)* 5, no. 1 (2009)
- Marti, Sergio, Thomas J. Giuli, Kevin Lai, and Mary Baker. "Mitigating routing misbehavior in mobile ad hoc networks." In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pp. 255-265.
- Djahel, Soufiane, Farid Nait-Abdesselam, and Zonghua Zhang. "Mitigating packet dropping problem in mobile ad hoc networks: proposals and challenges." *IEEE communications surveys & tutorials* 13, no. 4 (2011): 658-672.
- Nobahary, S., Garakani, H. G., Khademzadeh, A., & Rahmani, A. M. (2019). Selfish node detection based on hierarchical game theory in IoT. *EURASIP Journal on Wireless Communications and Networking*, 2019
- P.Manjula,Dr.S.Bagavathipriya" Intelligent Chimp Metaheuristics Optimization with Data Encryption Protocolfor WSN"Intelligent Automation and Soft Computing Vol 32 ,no.1(2022): 574-587 .
- Das, Semanti, and Abhijit Das. "An algorithm to detect malicious nodes in wireless sensor network using enhanced LEACH protocol." In *Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in*, pp. 875-881. IEEE, 2015.
- Meeran, A., Praveen A. N., & Ratheesh T. K. (2017). Enhanced system for selfish node revival based on watchdog mechanism. *2017 International Conference on Trends in Electronics and Informatics (ICETI)*.
- Varshney, T., Sharma, T., & Sharma, P. (2014). Implementation of Watchdog Protocol with AODV in Mobile Ad Hoc Network. *2014 Fourth International Conference on Communication Systems and Network Technologies*.
- Atakli, Idris M., Hongbing Hu, Yu Chen, Wei Shinn Ku, and Zhou Su. "Malicious node detection in wireless sensor networks using weighted trust evaluation." In *Proceedings of the 2008 Spring simulation multiconference*, pp. 836-843. Society for Computer Simulation International, 2008.
- Zhan, Guoxing, Weisong Shi, and Julia Deng. "Design and implementation of TARF: a trust-aware routing framework for WSNs." *IEEE Transactions on dependable and secure computing* 9, no. 2 (2012): 184-197.

An Improved Linguistic Haar Fuzzy Decision Maps

Karthik S

Department of Mathematics
Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology
Chennai, Tamil Nadu, India
karthik.kaka123@gmail.com

Gunasekar T

Department of Mathematics
Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology
Chennai, Tamil Nadu, India
tguna84@gmail.com

A Felix

Mathematics Division
School of Advanced Sciences
Vellore Institute of Technology
Chennai, Tamil Nadu, India
felix.a@vit.ac.in

Selvaraj A

Department of Mathematics
Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology
Chennai, Tamil Nadu, India
drselvaraja155@gmail.com

Abstract— *Fuzzy Decision Map (FDM) was introduced as an extension of DEMATEL in 2006. It integrates the salient features of eigenvalue method, fuzzy cognitive maps, and weighting equation. Usually, researchers facilitated triangular and trapezoidal fuzzy numbers to describe weights of causal relation between factors. But it is not always plausible to use triangular and trapezoidal number. Therefore, the objective of this research is to enhance Fuzzy Decision Map by using the linguistic terms to define causality among the factors. And, heptagonal fuzzy numbers are employed to represent the linguistic terms. Furthermore, the Haar ranking method has been proposed to reduce uncertainties. The proposed Improved Linguistic Haar Fuzzy Decision Maps (ILHFDM) has been applied to analyze the environmental benefits of organic farming. After applying the proposed model, it is identified that ecological services ranked as the most influencing factor among the rest of the factors in enhancing the environment by practicing organic farming.*

Keywords— Linguistic Variable, Fuzzy Number, Heptagonal Fuzzy Number, Haar wavelet, Fuzzy Decision Map.

I. INTRODUCTION

Decision making is an important activity in human beings' everyday life. Taking a decision in selecting the given alternatives is tedious when the influencing factors are increasing. Hence, researchers have proposed and used AHP, ANP, DEMATEL, FDM and FCM to tackle MCDA problems. Decision-making problems always concern about selecting the better alternatives [31]. In classical MCDM problems, weights between criteria were expressed by crisp values. Hence, it is easy for ranking alternatives and selecting the best criteria without any effort. In real world decision environment, classical MCDM faces complex problems in evaluating alternatives due to inherent uncertainties in the information. To tackle these kinds of situations, fuzzy logic [32, 33] is incorporated in MCDM where each causal relationship between alternatives will be given in terms of linguistic values [14]. Linguistic values are represented by the fuzzy number [15]. With this notion, more MCDA models like Fuzzy TOPSIS, Fuzzy ELECTRE, Fuzzy DEMATEL, Fuzzy VIKOR, Fuzzy DM etc., were proposed [25, 29].

Yu and Tzeng [31] introduced the concept of Fuzzy Decision Map (FDM) in 2006 as an extension of DEMATEL method. FDM integrates eigenvalue method, FCM, and weighting equation [27]. Fuzzy Cognitive Maps (Kosko, 1986) is an extension of Cognitive Maps (Axelord, 1974) using graphs. Each node in the graph describes the attributes and the link (edge) between the nodes describes the relationship between attributes. FCM also uses matrix method where each entry of the matrix describes the causal relationship between attributes. The FDM method has been efficaciously useful in different fields such as medical diagnosis [1, 2], land management, environment management [16], auto mobile industries [11], etc., Hence, implementing FDM in uncertain environment is a hard-core task. Therefore, this present study focuses on improving the FDM to Haar Wavelet concept using linguistic values. Haar transform is one of the simplest wavelets transform which is being used in image and signal compression in computer and electrical engineering as it is analyzing the local aspects of a signal in efficient way.

Due to over population and industrialization, people have shifted to conventional agriculture from organic agriculture which uses chemical fertilizers, pesticides, GMOs and additives for livestock which affects the health of the earth and other living organisms including human kinds in many ways. Only 1% of total agricultural lands are used for organic agriculture worldwide. Organic farming is a method of producing agricultural products without using chemical fertilizers, pesticides, GMOs and other additives for livestock. Organic farming focuses on sustainability and protects ecosystem to continue the life cycle on the earth and it helps to overcome the climate change. Instead of chemical fertilizers, manures and composts are used in organic farming which enrich the soil fertility. Organic farming is increasing rapidly worldwide due to its abundant benefits to all kinds of organisms. Organic farming is eco-friendly which enhances soil fertility, water quality, reducing climate change, etc., There are numerous environmental benefits in practicing organic farming such as Soil, Sustainability over the long term, Water, Biodiversity, Air and Climate Change, reduction of Genetically Modified Organisms (GMO), Ecological Services reported by Food and Agriculture Organization (FAO) of the United Nations (UN) [10].

Hence, this present study explores the environmental benefits of organic farming to create awareness among the farmers. To analyze this issue, Linguistic Haar Fuzzy Decision Maps (LHFDM) is proposed by integrating FDM and Haar wavelet technique. This paper is structured into five sections. Review of Literature on FCM and FDM is given in Section two. Preliminaries are given in section three to support the proposed method. In section four, Linguistic Haar Fuzzy Decision Maps is proposed. To illustrate the proposed method, the environmental benefits of organic farming are analyzed in section five and final section derives the conclusions and future directions.

II. REVIEW OF LITERATURE

Fuzzy Cognitive maps are widely applied in lot of domains such as Medical Diagnosis, Agriculture, Soil prediction, Engineering, Psychology, Social science etc., In 2009, Papageorgiou, Markinos and Gemtos used fuzzy cognitive maps to investigate the process of predicting cotton production using soft computing and predicted the yield of cotton production by the cause-and-effect method between soil properties and cotton yield [23]. In 2012 Zhang, et.al., used fuzzy cognitive mapping to develop environmental management of coal-mine ecosystems [34]. In 2013, fuzzy DEMATEL is one of the decision maps models, which has been used in analyzing the behavior of the youth by Devadoss and Felix [5, 6]. In 2013, Papageorgiou et.al., formulated FCM for yield prediction in apples with the aid of experts' opinion [22]. In 2014, Vasslides and Jensen analyzed the complicated links between ecological and social notions using FCM in highly intricate ecosystem [28]. In 2015, Mbugwa, Prager and Krall conducted a case study which integrated Geographic Information Systems and Spatial Decision Support Systems for the farmers and decision makers to enhance agriculture productivity by ranking the land suitability for agriculture using fuzzy logic [20]. In 2015, Jayashree et.al., have developed a FCM model and conducted a case study to predict the yield of coconut in Kerala [12]. In 2017, Bevilacqua et.al., utilized FCM mechanism to explore clinical risks of drug management in hospitals [4]. In 2017, Lamata et.al., contributed to determine optimal solution using fuzzy decision-making model [17]. In 2017, Joaquín et.al., implemented FCM to develop a fluvial system and applied this method for the case study of a Mediterranean fluvial management. This method can be used for forecasting river ecosystem [9]. In 2018, Karmarkar and Gilke developed a model based on fuzzy logic for assisting manufacturer by considering the customer needs and picking suitable automobile components that satisfies customers' need [13].

Nowadays, researchers are interested in constructing the hybrid techniques by combining two or more efficient method, which gives greater advantage than the individual model. In 2012 Liu et.al., constructed genetic optimized fuzzy inference model to evaluate suitability of agricultural land by integrating fuzzy inference system [18]. In 2014 Massei et.al., proposed the integration of MCDA and GIS for land management problems [19]. In 2014, Tadic et.al., framed an innovative hybrid MCDM model which combined fuzzy ANP, fuzzy DEMATEL and fuzzy VIKOR for selecting the city logistics concepts. This method can also be

useful for all kinds of MCDM problems in fuzzy environment [26].

In 2015, Baykasoglu and Gölcük developed a novel model to solve MADM problems in fuzzy environment which integrated FCM and hierarchical fuzzy TOPSIS. The main characteristic of this model is its capability to handle interdependencies between the problem attributes and ability of managing uncertainties under fuzzy environment [3]. In 2015, Natarajan, Subramanian and Papageorgiou proposed a hybrid approach using FCM for sugarcane yield classification and compared with the experimental results for the accuracy of the model [21]. In 2017, Wu, Liu and Chi proposed wavelet fuzzy cognitive maps which are combined by both wavelet theory and fuzzy cognitive maps. This model is more effective for the experiment on pattern classification problems and synthetic data and this can be used for even more complex systems [30]. In 2017, Rodríguez, Maradei and Escalante developed an integrated Geographic Information Systems based on Fuzzy AHP. This model helped to find the most suitable combinations of alternative land for lowest environmental impact for the bio energy crops [24].

From the review, it is observed that the research can be made in the following. (i) Linguistic Haar Fuzzy Decision Maps can be constructed using heptagonal fuzzy number. (ii) Wavelet technique for heptagonal fuzzy number may be proposed to reduce the uncertainty in Fuzzy Decision Maps and (iii) to illustrate this technique, environmental benefits of organic farming may be analyzed.

III. THE PROPOSED IMPROVED LINGUISTIC HEPTAGONAL FUZZY DECISION MAPS

This ILHFDM can make use of linguistic values for crisp values. So, the proposed ILHFDM tackles the practical ambiguity generally arises in the real-world problems in a proficient manner than the FDM model.

Now, the following steps are needed to execute the proposed ILHFDM method

Step 1: Generate pair wise comparison decision matrix between the criteria of the decision-making system with the field experts. HFNs are utilized in pair wise comparison matrix to convey the expert preferences in terms of linguistics.

Let the preference matrix be $\tilde{H} = [\tilde{h}_{ij}]$ where \tilde{H} represents $n \times n$ matrix and \tilde{h}_{ij} is the significance of criterion C_i with respect to criterion C_j using the fuzzy preference scale.

After preparing the heptagonal fuzzy pair wise judgment matrix, derive the fuzzy local weight vector (\tilde{U}) by using approximate fuzzy eigenvalue method as follows:

- Transform the heptagonal fuzzy pair wise judgment matrix into Haar heptagonal fuzzy pair wise comparison matrix using Haar wavelet technique.
- Add up every row of Haar fuzzy pair wise comparison matrix ($Ha_{\tilde{P}}$).

$$Sum(Ha_{\tilde{P}}) = (\delta_1 + \delta_2 + \eta_1 + \eta_2 + \beta_1 + \beta_2 + \beta_3 + \beta_4) \quad (1)$$

- Add every row of the summed Harr heptagonal

fuzzy pair wise comparison matrix to determine local weight vector.

$$\tilde{U} = [\tilde{U}_1, \tilde{U}_2, \dots, \tilde{U}_n]^T \quad (2)$$

- Choose the maximum element \tilde{r} in \tilde{U} for the normalization and normalize \tilde{U} by dividing \tilde{r} in each entry.

$$\tilde{r} = \max(\tilde{U}) \quad (3)$$

$$\tilde{U}_n = \frac{1}{\tilde{r}} \tilde{U} \quad (4)$$

Where \tilde{U}_n is the normalized fuzzy local weight vector.

Step 2: Influence among criteria can be determined by

$$\tilde{C}^{k+1} = \tilde{G}(\tilde{C}^k \cdot \tilde{W}), \tilde{C}^0 = \tilde{I}_{n \times n} \quad (5)$$

Where $\tilde{I}_{n \times n}$ represents the identity matrix of order n , $\tilde{W} = (\tilde{w}_{ij})$ is a $n \times n$ weight matrix with HFNs, that collects the values of connecting edge weight between concepts C_i and C_j , $\tilde{C}^{(k+1)}$ and $\tilde{C}^{(k)}$ are the state matrices at iterations $(k+1)$ and (k) , respectively, $\tilde{C}^{(0)}$ is the initial matrix with and \tilde{G} is a threshold transformation function.

Here, the following hyperbolic-tangent function has been taken as threshold function.

$$\tilde{G}(x) = \left(\frac{1-e^{-x}}{1+e^{-x}} \right) \quad (6)$$

Now, FCM is designed with the help of experts to specify the influence between criteria. Then, Haar wavelet techniques is applied to all the entries of the heptagonal FCM matrix and sum each element in a Haar heptagonal FCM matrix to get the summed Haar heptagonal FCM matrix. Hence, the FCM model is applied using Eq. (5) till it reaches steady-state (\tilde{C}).

Step 3:

Fuzzy global weight \tilde{X} can be determined by

$$\tilde{X} = \tilde{U}_n + \tilde{C}_n \tilde{U}_n \quad (7)$$

Where \tilde{U}_n is the normalized fuzzy local weight vector and \tilde{C}_n is the normalized fuzzy steady-state matrix.

- Choose maximum value $\tilde{\mu}$ of the fuzzy global weight vector \tilde{X}

$$\tilde{\mu} = \max(\tilde{X}) \quad (8)$$

- Normalization is done by

$$\tilde{X}_n = \frac{1}{\tilde{\mu}} \tilde{X} \quad (9)$$

Where \tilde{X}_n is the required normalized fuzzy global weight vector. Finally, ranking is given for all the factors with respect to the values of \tilde{X}_n and the ranking follows the descending order. That is, maximum value of \tilde{X}_n gets first rank and remaining ranks are given according to their values.

IV. ILLUSTRATION TO THE PROPOSED ILHFDM

Now a days, farmers in all over the world turn into organic agriculture rather than continuing conventional agriculture. Although earning more money by conventional agriculture, farmers are now aware of organic farming due to its fabulous effect on both human health and ecosystem. Organic farming maintains the life cycle without harming the ecosystem which is essential for all biodiversity to survive on the earth and any changes in the life cycle led to enormous impact on ecosystem. In order to make the earth more sustainable environmentally, organic farming is essential. So, this paper is aimed to study the cause and effect of each factor which enhances environmental sustainability through organic farming and ranking is determined according to environmental benefits of organic farming using ILHFDM. The below mentioned factors are influencing each other one in many ways in enhancing the benefits of environment [9] that are difficult to handle in real life. Therefore, fuzzy logic is applied to the problem where linguistic values are employed to describe the causal relationship between factors.

A₁- Sustainability over the long term: Conversion to organic farming provides healthy food, waste reduction, quality water, cost savings, more profit, etc., without damaging natural resources. Organic farming is a factor of sustainable development for society, environment, and economy over the long term.

A₂- Soil: In organic farming, some techniques are carried out such as minimum tillage, intercropping, cover crops, crop rotations, organic fertilizers, manures, composts, etc., to enhance the soil fertility and to reduce the soil erosion. Soil building is one of the vital parts in organic farming because soil contains nutrients which help to maintain the flora and fauna of the earth.

A₃- Water: Pollution of ground water is related to usage of synthetic fertilizers and pesticides in the agricultural field where as organic farming does not contribute to ground water pollution, the system does not entertain to use those kinds of chemical fertilizers and pesticides in the field. Also, organic farming consumes less water than commercial farming system.

A₄- Air and Climate Change: Organic farming diminishes greenhouse gas emissions through carbon sequestration by the soil.

A₅- Biodiversity: Practicing organic farming sustains diverse combination of plants, micro and macro-organisms to optimize nutrients and energy cycling for the agricultural production. Organic farming feeds worms are microbes in the soil which are essential for the crop's growth. The lack of usage of chemical fertilizers and pesticides produce more biodiversity in the surroundings.

A₆- Reduction of Genetically Modified Organisms (GMO): Genetically Modified Organisms are the result of DNA alteration from another organism through genetic engineering

techniques. Genetically Modified Organisms are not allowed in any stages of organic farming because of their impact on environment and human health. Recent study shows that GMO contained foods are associated with an increased risk of cancer.

A₇- Ecological Services: Ecological services are various benefits from natural resources that are freely accessible to human beings including all kinds of species on the earth. Organic farming plays a crucial role in improving ecological services as it reduces the degradation of natural resources and global warming.

The proposed ILHFDM model is examined using the above-described problem where a decision maker tries to choose the best influencing factor among seven criteria.

V. RESULTS AND DISCUSSION

The proposed LHFDM technique is applied to study environmental benefits of organic farming. The proposed LHFDM technique integrates eigenvalue method to derive fuzzy local vectors, FCM with linguistic variables to describe the influence among factors and fuzzy weighting equation. At first, fuzzy local weight vector is computed from the pair wise comparison matrix among criteria whose entries are heptagonal fuzzy numbers and heptagonal fuzzy reciprocal numbers. After that, Haar wavelet technique is applied to reduce the uncertainty. Then, the influences among criteria are determined using heptagonal FCM matrix and equation (5). Here, hyperbolic tangent function is assumed as threshold transition function. Then, fuzzy steady state matrix is obtained after three iterations. Finally, fuzzy global weight vector is computed. After that, ranking has been given to all factors based on normalized global weight vector.

Table 1: Ranking of factors

Factors	Global weight vector	Normalized vector	Rank
A ₁	1.744719	0.998157	3
A ₂	1.74234	0.996796	5
A ₃	1.739378	0.995102	7
A ₄	1.74571	0.998724	2
A ₅	1.740364	0.995666	6
A ₆	1.742511	0.996894	4
A ₇	1.74794	1	1

From the Table 1, it is observed that A₇- Ecological Services gets first rank as it has highest normalized global weight vector among other factors. Even all the factors are influenced by practicing organic farming, ecological services is the most influencing factor among them. The advantage of this model over existing model given in [8] is

that the proposed model takes less iterations to attain steady state matrix and is effective in handling linguistic terms and solving MCDM problems.

VI. CONCLUSION

This present study extends FDM technique to LHFDM technique to rank the environmental beneficial factors through organic farming. After ranking the factors, it is observed that each factor has almost equal importance in enhancing the environment. Even all the factors have almost equal importance, A₇- Ecological Services gets first as it has highest value among other factors. In future, this method can be applied to medical diagnosis, agricultural management, environmental management, energy management and so on. Moreover, this method can also be extended using the fuzzy numbers such as Octagonal, Nonagonal, and Decagonal for representing linguistic terms.

REFERENCES

- [1] Amirkhani, A., Papageorgiou, E.I., Mohseni, A., Mosavi, M. R., A review of fuzzy cognitive maps in medicine: Taxonomy, methods, and applications, Computer Methods and Programs in Biomedicine 142 (2017) 129- 145.
- [2] Amirkhani, A., Papageorgiou, E.I., Mosavi, M. R., Mohammadi, K., A novel medical decision support system based on fuzzy cognitive maps enhanced by intuitive and learning capabilities for modeling uncertainty, Applied Mathematics and Computation 337 (2018) 562- 582.
- [3] Baykasoglu, A., Gölcük, I., Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS, Information Sciences 301 (2015) 75- 98.
- [4] Bevilacqua, M., Ciarapica, F. E., Mazzuto, G., Fuzzy cognitive maps for adverse drug event risk management, Safety Science 102 (2018) 194-210.
- [5] Devadoss, V.A., Felix, A., A new Fuzzy DEMATEL method in an uncertain linguistic environment, Advances in Fuzzy Sets and Systems, Volume 16(2), 93-123, 2013.
- [6] Devadoss, V.A., Felix, A., Fuzzy DEMATEL approach to study cause and effect relationship of youth violence, International Journal of Computing Algorithm, 2, 363-372, 2013.
- [7] Dzitac, I., The fuzzification of classical structures: a general view, International Journal of Computers, Communications & Control 2015; 10(6): 772-788.
- [8] Elomda, B. M., Hefny, H. A., Hassan, H. A., An extension of fuzzy decision maps for multi-criteria decision-making, Egyptian Informatics Journal, 14(2), 147-155, 2013.
- [9] Gutiérrez, J. S., Rincón, G., Alonso, C., García-de-Jalón, D., Using fuzzy cognitive maps for predicting river management responses: A case study of the Esla River basin, Spain, Ecological Modelling 360 (2017) 260-269.
- [10] <http://www.fao.org/organicag/oa-faq/oa-faq6/en/>
- [11] Jassbi, J., Mohamadnejad, F., Nasrollahzadeh, H., A Fuzzy DEMATEL framework for modeling cause and effect relationships of strategy map, Expert Systems with Applications 38 (2011) 5967- 5973.
- [12] Jayashree, L. S., Palakkal, N., Papageorgiou, E.I., Papageorgiou, K., Application of fuzzy cognitive maps in precision agriculture: a case study on coconut yield management of southern India's Malabar region, Neural Comput & Applic 26(8), (2015) 1963-1978.
- [13] Karmarkar, U., Gilke, N. R., Fuzzy Logic based decision support systems in variant production, Materials Today: Proceedings 5 (2018) 3842-3850.
- [14] Karthik, S., Saroj Kumar Dash, Punithavelan, N., A Fuzzy Decision-Making System for the Impact of Pesticides Applied in Agricultural Fields on Human Health, International Journal of Fuzzy System Applications, Volume 9(3), pp42-62, 2020.
- [15] Karthik, S., Saroj Kumar Dash, Punithavelan, N., Haar Ranking of Linear and Non-Linear Heptagonal Fuzzy Number and Its

- [16] Application, International Journal of Innovative Technology and Exploring Engineering, Volume 8(6), pp1212-1220, 2019.
- [17] Kyriakarakos, G., Patlitzianas, K., Damasiotis, M., Papastefanakis, D., A fuzzy cognitive maps decision support system for renewables local planning, *Renewable and Sustainable Energy Reviews* 39 (2014) 209–222.
- [18] Lamata, M. T., Pelta, D., Verdegay, J. L., Optimisation problems as decision problems: The case of fuzzy optimisation problems, *Information Sciences* 000 (2017) 1–12.
- [19] Liu, Y., Jiao, L., Liu, Y., and Jianhua., A self-adapting fuzzy inference system for the evaluation of agricultural land, *Environmental Modelling & Software* 40 (2013) 226–234.
- [20] Massei, G., Rocchi, L., Paolotti, L., Greco, S., Boggia, A., Decision Support Systems for environmental management: A case study on wastewater from agriculture, *Journal of Environmental Management* 146 (2014) 491–504.
- [21] Mbugwa, G., Prager, S. D., Krall, J. M., Utilization of spatial decision support systems decision-making in dryland agriculture: A Tifton burclover case study, *Computers and Electronics in Agriculture* 118 (2015) 215–224.
- [22] Natarajan, R., Subramanian, J., Papageorgiou, E. I., Hybrid learning of fuzzy cognitive maps for sugarcane yield classification, *Computers and Electronics in Agriculture* 127 (2016) 147–157.
- [23] Papageorgiou, E. I., Aggelopoulou, K. D., Gemtos, T. A., Nanos, G. D., Yield prediction in apples using Fuzzy Cognitive Map learning approach, *Computers and Electronics in Agriculture* 91 (2013) 19–29.
- [24] Papageorgiou, E. I., Markinos, A. T., Gemtos, T. A., Fuzzy cognitive map-based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application, *Applied Soft Computing* 11 (2011) 3643–3657.
- [25] Rodríguez, R., Maradei, P. G., and Escalante, H., Fuzzy spatial decision tool to rank suitable sites for allocation of bioenergy plants based on crop residue, *Biomass and Bioenergy* 100 (2017) 17–30.
- [26] Shemshadi, A., Shirazi, H., Toreihi, M., Tarokh, M.J., 2011. A fuzzy VIKOR method for supplier selection based on entropy measure for objective weighting. *Expert Syst. Appl.* 38 (10), 12160- 12167.
- [27] Tadic, S., Zecevic, S., Krstic, M., A novel hybrid MCDM model based on fuzzy DEMATEL, fuzzy ANP and fuzzy VIKOR for city logistics concept selection, *Expert Systems with Applications* 41 (2014) 8112–8128.
- [28] Tzeng, G. H., Chen, W. H., Yu, R., Shih, M. L., Fuzzy decision maps: a generalization of the DEMATEL methods, *Soft Comput* (2010) 14:1141–1150.
- [29] Vasslides, J. M., and Jensen, O. P., Fuzzy cognitive mapping in support of integrated ecosystem assessments: Developing a shared conceptual model among stakeholders, *Journal of Environmental Management* 166 (2016) 348–356.
- [30] Wang, J. J., Jing, Y. Y., Zhang, C., Zhao, J. H., Review on multi-criteria decision analysis aid in sustainable energy decision-making, *Renewable and Sustainable Energy Reviews* 13 (2009) 2263–2278.
- [31] Wu, K., Liu, J., Chi, Y., Wavelet fuzzy cognitive maps, *Neurocomputing* 232 (2017) 94–103.
- [32] Yu, R., Tzeng, G. H., A soft computing method for multi-criteria decision making with dependence and feedback. *Applied Mathematics and Computation*, 180, 63–75, 2006.
- [33] Zadeh L. A., Fuzzy sets. *Information Control* 8(3):338–353, 1965.
- [34] Zadeh L. A., The concept of a linguistic variable and its application to approximate reasoning (Part II). *Information Science*, 8:301–357, 1975.
- [35] Zhang, H., Song, J., Su, C., Mengying, Human attitudes in environmental management: Fuzzy Cognitive Maps and policy option simulations analysis for a coal-mine ecosystem in China, *Journal of Environmental Management* 115 (2013) 227–234.

Classification of Pulmonary Emphysema using Deep Learning

R.Ramadoss¹

Department of ECE

SRM Institute of Science and Technology
Chennai, India

rr5606@srmist.edu.in

C.Vimala²

Department of ECE

SRM Institute of Science and Technology
Chennai, India

vimalac@srmist.edu.in

Abstract— Emphysema is one of the lung diseases that comprise COPD. Emphysema is a long-term lung illness that results in alveolar destruction, which destroys tiny air sacs in the lung that allow fast gas exchange or exchange of oxygen and carbon dioxide molecules between the bloodstream and the alveoli. It is one of the two most common kinds of COPD. Another significant COPD consequence is chronic bronchitis. Heart failure collapsed lungs, and big holes in the lungs are all possible consequences of emphysema. Emphysema is an incurable lung airway condition. This research chooses chest X-rays rather than CT scans. This article aims to create a deep learning model that can identify emphysema patients from chest radiographic images. This paper chooses 1696 chest X-ray images from an NIH Chest X-ray dataset. ResNet50 is used as a deep learning model since it is proved effective in medical image classification. This model is trained using multiple epochs and mini-batch sizes. It achieved 98.82 percent accuracy, 1.0 value precision for emphysema, 0.9885 for AUC, 97.70 percent recall/sensitivity for emphysema, 100 percent specificity, 98.85 percent balanced accuracy, and an 18.3303 geometric mean. Additionally, we obtained a false-positive rate of 0%.

Keywords— *Pulmonary Emphysema classification, Chronic Obstructive Pulmonary Disease, Deep Learning*

I. INTRODUCTION

Around 12 million individuals have been diagnosed with chronic obstructive pulmonary disease (COPD). In contrast, another 3.1 million have been diagnosed with emphysema, both of which cause lung tissue to weaken, resulting in difficulty breathing [1]. Emphysema is associated with COPD. However, not everyone who has COPD gets emphysema. Emphysema is a lung disorder under chronic obstructive pulmonary disease (COPD) [2]. Normal lung tissue resembles a swab, but emphysema-affected lung tissue resembles a deteriorated porous swab with a limited ability to "spring back" into shape [3]. Emphysema is a continuous ailment that often begins in tiny gaps among the alveoli, gradually close to generating more significant air gaps [3]. In these air spaces, old air is trapped, making it harder for individuals to breathe new air. Emphysema patients struggle to live and get adequate oxygen due to reduced blood flow via the alveoli. Pulmonary emphysema occurs when the pulmonary alveoli are damaged due to prolonged smoking. Once collapsed, pulmonary alveoli never rebuild themselves [5]. That implies that early identification of pulmonary emphysema is critical for delaying the disease's development. Doctors diagnose pulmonary emphysema in two ways. Spirometry and diagnostic imaging [5]. Lung disease has traditionally been discovered via blood tests, sputum sample tests, chest X-ray examinations, and computed tomography (CT) scan examinations [6]. Machine learning and deep learning are two subgroups of artificial intelligence, where a model incorporates data to classify

models directly from images, text, or sound. [7] Typically, deep learning has achieved via a neural network design. The word "deep" refers to the network's layer count; the more layers, the more complex the network. While conventional neural networks have just two or three layers, deep neural networks may have more than three [8], [9]. This paper will discuss diagnosing emphysema using a deep learning network, namely ResNet-50. This paper analyzes the test's accuracy, balanced accuracy, Geometric mean, ROC, AUC, precision, sensitivity, and specificity over several training epochs.

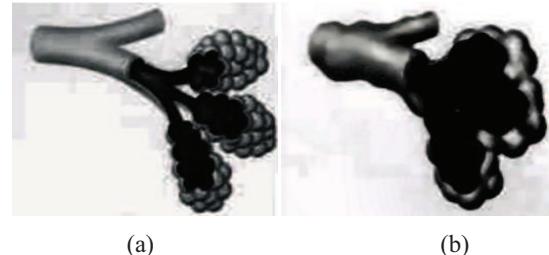


Fig. 1. (a) Normal alveoli (b) Collapsed alveoli [10]

II. RELATED WORK

Numerous studies are published in literature employing deep learning models to classify pulmonary emphysema using chest X-ray or CT scan images. When it comes to classifying the emphysema subclass of COPD, a novel learning approach based on probabilistic PCA is demonstrated for lung CT sample classification. The purpose of this study is to build a computer-assisted approach that makes use of an enhanced deep learning strategy. [5] The lung segmentation is performed using the FCM approach. Additionally, the ALTP algorithm created the pattern picture from the segmented image patterns. Using suggested PO-FRCNN with the ALTP picture output as input classification is performed. The use of IRDA through ALTP and PO-FRCNN to image processing in diagnosing emphysema resulted in successful pattern extraction and classification of the illness into four distinct groups. When contrasted with modern approaches, this suggested method produces the most outstanding results and is most successful in diagnosing emphysema. The experimentation is conducted using benchmark and real-time datasets.

Deep learning has been used in more and more medical imaging procedures in the last few years.[8] The goal of this paper is to develop a classifier for GOLD classification using CT images that have been trained with and without preprocessing. The method used in this research is based on the ResNet-50 network. The results in this paper achieve 0.8 of AUC.

This study describes a CNN-based sequence of COPD, and respiratory disease screening predicts the coherence of cellular breakdown inside the lungs (NLST) [9]. Spirometric COPD and visual emphysema from COPD gene research are utilized in this work. Finally, it achieves an accuracy of 79.8%.

This paper is based on the classification of emphysema in lung CT scans. This paper is focused on classifying the ailment using lung computer tomography images. [12] The methods include multi-scale deep CNNs such as residual networks with 20 layers to extract more high-level information. The results have achieved 92.68% accuracy on 91 HRCT datasets.

Emphysema is a medically serious lung disease that results from a series of smoking-related lung disorders.[13] Noise filtering was used on CT scans as part of low-dose lung cancer screening methods to improve the reliability of pulmonary compactness rates for long-term study.

To explore if the emphysema pattern at the participant level could be used to predict disability and death when classified using a deep learning algorithm[14]. Centrilobular, Panlobular, and Paraseptal emphysemas, often known as CLE, PLE, and PSE, are the three kinds of pulmonary emphysema [15]. Here, they suggest an unsupervised method to discover lung macroscopic patterns (LMPs) resultant of lung areas, which predetermine the emphysematous zones using a variation of the Latent Dirichlet Allocation (LDA) replication [15]. This system can identify highly repeatable LMPs that accurately predict the standard emphysema subtypes. CNN's architecture represented a significant advancement in medical image analysis. [16] Until 2017, various promising architectures demonstrated their efficiency, beginning with AlexNet in 2012 and progressing to Inception-**V4 in 2017, going via VGG, GoogLeNet, DenseNet, and ResNet. As a result of this thorough research, it became clear that using natural images as training datasets might be beneficial in a variety of applications, including medical image analysis.

The work in this paper is focused on pulmonary Emphysema Quantitative Analysis using Local Binary Patterns [17]. The method used in this research is LBP, which is used to represent texture features. This paper achieves 95.2% of accuracy on 168 2D HRCT datasets.

Emphysematous lesions have fewer alveolar septa per unit volume [18]. This article focuses on assessing the geographic extent of the observed low-attenuation zones, applying the Euclidean distance transformation to investigate the relationship between emphysematous lesions and bronchial or pneumonic veins. These procedures might be used to effectively eliminate emphysematous damage and could be used to quantify their design of dispersion. CT scans from five different instances were utilized in this study (one normal and four abnormal). [24] The approach employed in this study is CAD-based 3-D pulmonary imaging with emphysematous lesions. Finally, a 3-D area growth approach was used to collect and analyze emphysematous lesions.

Several strategies and algorithms from various articles were discovered as a result of the aforementioned study. The detection of emphysema may be accomplished using many approaches. Many of the authors in the studies listed the above-employed machine learning methods. As a result, we

proposed a deep learning model ResNet-50 using chest x-ray to evaluate precision, AUC, Sensitivity, specificity, balanced accuracy, geometric mean, ROC (TPR vs. FPR) using confusion matrix in this paper.

III. METHODOLOGY

This section will begin by discussing the dataset used in this work. The proposed model of emphysema classification consists of chest X-ray datasets prepared from two publicly available sources. [18-20] One is an arbitrary NIH Chest X-ray Dataset, and the other is an NIH Chest X-ray 8 Dataset. In total, 1696 (1024 x 1024) png format frontal-view X-ray images are in the database, of which 848 contain Emphysema. The remaining 848 images are Normal cases. In this research, we make use of 340 images taken for testing purposes, whereas 1356 images are obtained for training purposes.

Diagnostic and detection technologies provide a helpful second opinion to physicians and aid in screening. This approach would also aid in the rapid delivery of data to physicians. This article uses MATLAB to apply a Deep Learning (DL) method to identify emphysema on chest radiographs. In deep learning, feature extraction is not required. The proposed system includes the following stages: Initial stage the images were gathered from a publicly accessible database. The preprocessing step of the image analysis process is essential. They are preprocessing aims to remove unnecessary details from images and improve image quality. Automatic preprocessing will occur throughout training. This preprocessing is necessary because pictures might be grayscale or color, which requires a single matrix image for grayscale images and three matrices for color images. We need the exact dimensions of matrices during training, or else it will raise an error. After preprocessing, the preprocessed image is used to perform feature extraction and classification. We provide a pre-trained deep learning network, ResNet-50, for the classification of emphysema. ResNet-50 network trained on the Image Net data set. Why did we choose an x-ray dataset? The whole X-ray process typically takes less than 15 minutes. Because X-ray images are digital, they may be seen on a computer screen by a doctor within minutes.

A. Chest X-ray findings

In Fig 2, the chest x-ray PA view demonstrates bilateral emphysema alterations. Now let us examine the changes. The first is an increase in the lucency of the lung field. Increased lucency indicates that the lung area is darker than usual. Secondly, as compared to a typical chest X-ray, the rib gaps are expanded in this specific chest X-ray.

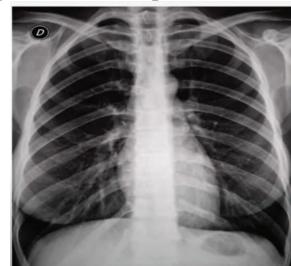


Fig. 2. Chest X-ray PA(Posterior Anterior) view

Tubular cardiac shadow is the third observation while flattening of diaphragm domes is the fourth finding. Essentially, this X-ray shows the four primary changes that indicate emphysema.

B. ResNet-50

ResNet is the abbreviation for the residual network. Visuals, auto-encoding, and classification are all used in combination with a deep CNN. [7], [8], [16] ResNet50 was refined on a subset of the ImageNet archive. The ResNet-50 flowchart and architecture are shown in Fig.3 & 4. It has convolution, max-pooling, and fully connected layers [22]. The first 174 linked layers were frozen, and the final connected layers were replaced (175 to 177). The trained network does not retain the properties of the updated frozen layers. Here, we may adjust the last three layers of our proposed model, namely the fully connected layer, the softmax layer, and the classification type, as shown in Fig. 5. The ResNet model and network layer graph are analyzed using the Deep Learning Network Analyzer.

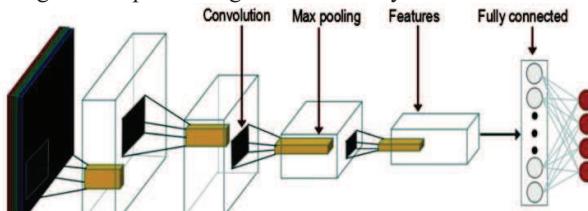


Fig. 3. Flowchart of ResNet50 [7]

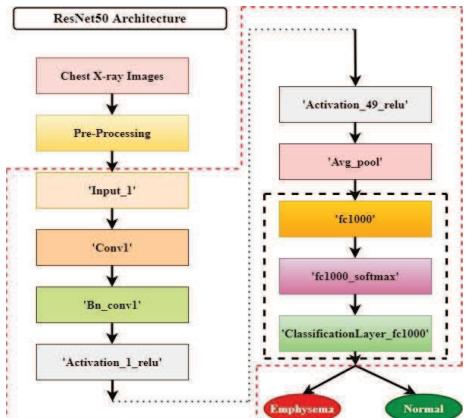


Fig. 4. The architecture of ResNet50

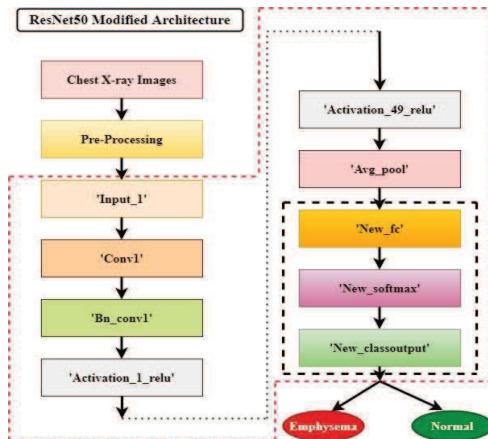


Fig. 5. ResNet50 modified architecture

During training, we use Adam optimizer include learning rate information, L_2 regularization factor ($1.0000e^{-4}$), Max epochs (set 5,8,10,20,30) and mini-batch size (set 6,20,32,128). The performance of this model was ascertained by precision, AUC, Sensitivity, Specificity, and ROC.

IV. RESULTS AND DISCUSSION

Our ResNet-50 with 177 layers and 192 connections gets 98.8235% test accuracy in 20 epochs of training. Here actual label we are using is 340 (170+170). In total 1696 images can be used here. Overall 340 images are used for testing purposes and 1356 for training purposes. Training option consists of maximum epochs is 20, Mini batch size 32, total sample emphysema 848, and Normal case 840.

A. Epochs

One forward and one reverse pass through all of the training examples. Each time we use the entire training dataset that is called an epoch. So, in our scenario, using Adam optimizer, we would have performed 10, 42, 67, 226 passes before finishing one epoch. Of the passes mentioned above, 42 iterations per epoch achieved good accuracy.

B. Mini-Batch Size

A Mini-Batch represents a subset of the training set to assess the gradient of loss functionality and weight updating. In this case, it has experimented with various mini-batch sizes, including 6, 20, 32, and 128. Out of these 32, it has achieved good accuracy.

C. ROC Analysis

$$TPR = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})} \quad (1)$$

$$FPR = \frac{\text{False Positives}}{(\text{False Positives} + \text{True Negatives})} \quad (2)$$

The TPR, also known as Sensitivity, is computed using the equation (1). The probability that an actual positive will test positive is called the TPR.

The FPR is based on the probability of rejecting the null hypothesis incorrectly. Simply put, while the fact is negative, the test predicts a favorable outcome. On the other hand, FPR is calculated using equation (2).

In our scenario, we experiment with various epochs and mini-batch sizes, including (5, 20), (8, 6), (10, 32), (10, 128), (20, 32), and (30, 32) as shown in Table I. Finally, with a maximum epoch of 20 and a mini-batch size of 32, we get a 0% false-positive rate as shown in Fig.6.

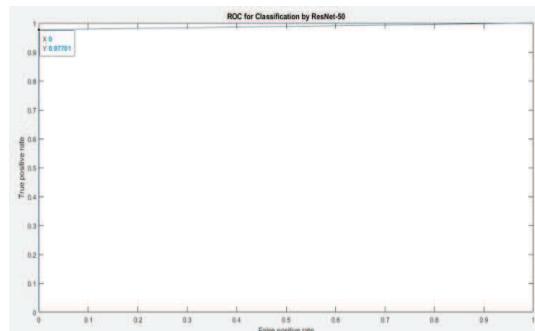


Fig. 6. ROC Curve analysis (20 Epochs and mini-batch size of 32)

TABLE I. ROC ANALYSIS

(Max Epochs, Mini Batch Size)	ROC	
	TPR	FPR
10, 128	0.9647	0.0353
5, 20	0.9767	0.0119
10, 32	0.9825	0.0118
20, 32	0.9770	0
30, 32	0.9873	0.0769
8,6	0.9821	0.0291

D. AUC

$$AUC = \int_0^1 ROC(t)dt \quad (3)$$

In summary, the ROC curve is an effective tool for determining the appropriate cut-off point for a specific screening method.

TABLE II. AUC CLASSIFICATION

RANGE	CLASSIFICATION
0.6 < AUC < 0.7	Not good
0.7 < AUC < 0.8	Worthless
0.8 < AUC < 0.9	Good
0.9 < AUC < 1.0	Excellent

Based on the AUC classification given from Table II the Performance Metrics are obtained with an excellent AUC range with various epochs and mini-batch sizes as shown in Table III.

E. Precision

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \quad (4)$$

F. Accuracy

$$Accuracy = \frac{\text{Sum of correct classification}}{\text{Total number of classification}} \quad (5)$$

G. Recall

A recall is also called Sensitivity or True positive rate. The main objective is to reduce false Negatives.

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \quad (6)$$

The result is a value that ranges between 0.0 and 1.0, with 0.0 denoting no recall and 1.0 denoting absolute or ideal recall.

H. Specificity

Specificity (SP) is denoted as the proportion of true negative forecasting to the total number of negative predictions. Additionally, it is referred to as true negative rate (TNR).

$$Specificity = \frac{True\ Negatives}{(True\ Negatives + False\ Positives)} \quad (7)$$

I. Geometric mean

The geometric mean of the TPR along with the TNR is given below.

$$Geometric\ mean = \sqrt{True\ Positives * True\ negatives} \quad (8)$$

J. Balanced accuracy

A balanced Accuracy is a classification approach that may be used for binary or multi-class classification. It is the arithmetic mean of sensitivity and specificity; it is useful

when dealing with data that is unbalanced, i.e., when one of the target groups is found much more often than the other.

$$Balanced\ accuracy = \frac{Sensitivity + Specificity}{2} \quad (9)$$

The outcomes are executed based on AUC Classification and the proposed work is summarized in Tables III and IV.

TABLE III. PERFORMANCE METRICS

(Max Epochs, Mini Batch Size)	ResNet50				
	Precision	AUC	Accuracy	Sensitivity	Specificity
10, 128	0.9647	0.9647	96.47%	96.47%	96.47%
5, 20	0.9882	0.9824	98.24%	97.67%	98.81%
10, 32	0.9882	0.9853	98.53%	98.25%	98.82%
20, 32	1.0	0.9885	98.82%	97.70%	100%
30, 32	0.9176	0.9552	95.29%	98.73%	92.31%
8,6	0.9706	0.9765	97.65%	98.21%	97.09%

TABLE IV.BALANCED ACCURACY VS GEOMETRIC MEAN

(Max Epochs, Mini Batch Size)	Balanced Accuracy	Geometric Mean
10, 128	0.9647	18.1108
5, 20	0.9824	18.2757
10, 32	0.9853	18.3030
20, 32	0.9885	18.3303
30, 32	0.9552	18
8,6	0.9765	18.2209

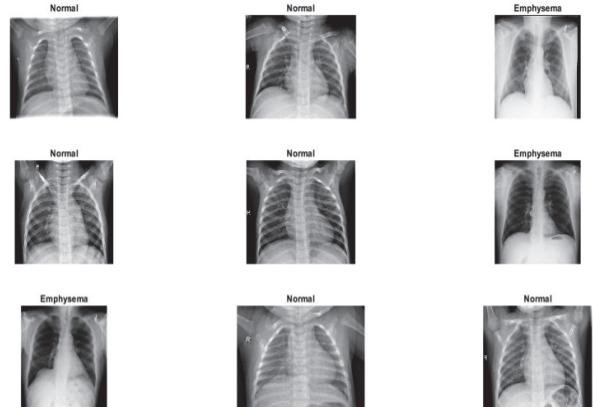


Fig. 7. Nine images were chosen at random based on their class.

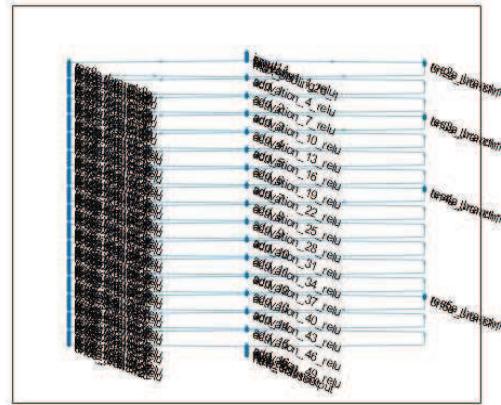


Fig. 8. Igraph (Extract all layers) (It looks dense because of 177 layers)

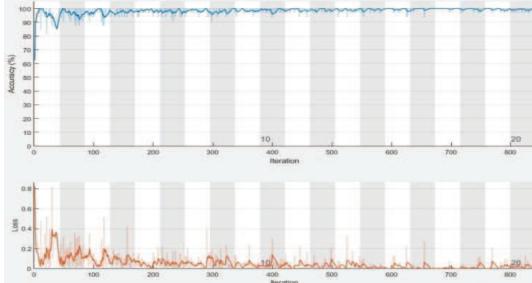


Fig. 9. The plot of Training accuracy and its losses

The performance of a classification algorithm is shown and described using a confusion matrix in Fig.10, which evaluates our algorithm's performance based on the testing dataset. Additionally, this matrix indicates which images were successfully classified and incorrectly classified. By examining the confusion matrix, we can monitor the performance of our deep learning model. We will analyze various performance metrics such as TP, TN, FP, FN, Sensitivity, Specificity, AUC, Accuracy, etc. Finally, we will express the results using several techniques for classifying emphysema and demonstrate that our approach is more accurate than previous studies, as shown in Table V.

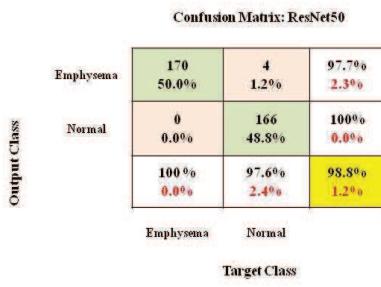


Fig. 10. Confusion Matrix

Using a Linear rectified unit from the proposed ResNet50 model as shown in Fig.11, displays dark colors that indicate less active features, whereas bright colors indicate strongly activated features. By randomly choosing an image from test data storage, classify the correct image as shown in Fig.12. the test score is 0.99907 is achieved.

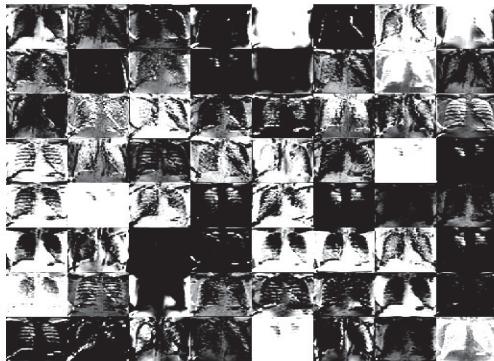


Fig.11. Activation of a Linear rectified unit (Montage)

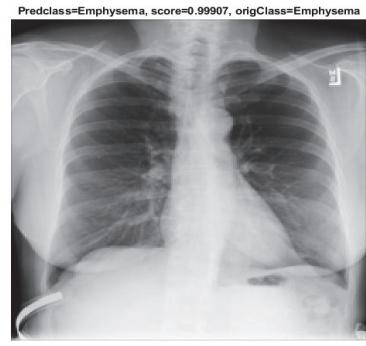


Fig.12. Classification of the Test dataset

TABLE V. Existing system versus proposed Resnet50 model

Author Details	Methodology	Accuracy
L. Peng et al. [12]	Multi-Scale Residual Network	92.68%
Sørensen, L., et al.[18]	LBP, Nearest Neighbor classifier	95.2%
Tuba, E., et al.[19]	SVM classifier with Elephant Herding Optimization	89.72%
Er, O., et al. [20]	LM Algorithm	95.43%
Isaac, A., et al. [21]	SVM-based Extreme Machine Learning classifier	(89.02% & 91.89%)
Bhuma, C. M.[22]	ImageNet	95.88%
Altan, G., et al. [23]	CNN	93.67%
Riti, Y.F. et al.[24]	MLP classifier	85%
Proposed System	Proposed Resnet50 model	98.82%

V. CONCLUSION AND FUTURE WORK

The proposed research is presented in a simple deep learning-based classification methodology for automated classification of Emphysema. It is the ultimate aim to avoid a false-positive rate, and our model achieves a zero false-positive rate. The ResNet50 CNN deep learning model tests the network with 177 layers, and 192 connections functioned optimally with an enhanced comprehensive precision along with AUC. This benchmark provides a complete picture of the performance regarding accuracy and recall. The overall accuracy is an excellent pointer as the evaluation dataset implemented in this article is consistently disseminated. This model is trained using multiple Epoch and Mini batch sizes. The proposed model has achieved 98.82 percent accuracy, 1.0 value precision, 0.9885 AUC, 97.70 percent recall/sensitivity for emphysema, 100 percent specificity, 98.85 percent balanced accuracy, the geometric mean of 18.3303 for emphysema, and the predicted score of 0.99907, additionally, a false positive rate of zero is also achieved. So this model will assist the medical professionals in Emphysema detection and deriving necessary verification, adding to it the testing data contains just four incorrectly classified images. This approach has accurately classified emphysema images in chest X-rays, once emphysema is diagnosed, the appropriate medication or surgery may be performed promptly to save mortality. In the future, this work can be extended to detect and classify all

types of pulmonary disease based on other deep learning models.

REFERENCES

- [1] www.emphysemafoundation.org
- [2] M. Prasad, A. Sowmya, and P. Wilson, "Multi-level classification of emphysema in HRCT lung images," *Pattern Anal. Appl.*, vol. 12, no. 1, pp. 9–20, 2009, doi: 10.1007/s10044-007-0093-7.
- [3] I. P. Peña *et al.*, "Automatic emphysema detection using weakly labeled HRCT lung images," *PLoS One*, vol. 13, no. 10, pp. 1–16, 2018, doi: 10.1371/journal.pone.0205397.
- [4] T. K. Liang, T. Tanaka, H. Nakamura, and A. Ishizaka, "A neural network based computer-aided diagnosis of emphysema using CT lung images," *Proc. SICE Annu. Conf.*, pp. 703–709, 2007, doi: 10.1109/SICE.2007.4421073.
- [5] W. L. Foster *et al.*, "The emphysemas: radiologic-pathologic correlations," *Radiographics*, vol. 13, no. 2, pp. 311–328, 1993, doi: 10.1148/radiographics.13.2.8460222.
- [6] S. Mondal, A. K. Sadhu, and P. K. Dutta, "Adaptive Local Ternary Pattern on Parameter Optimized-Faster Region Convolutional Neural Network for Pulmonary Emphysema Diagnosis," *IEEE Access*, vol. 9, pp. 114135–114152, 2021, doi: 10.1109/access.2021.3105114.
- [7] M. El-Gohary, J. F. De La Renaud Faverie, N. Gharbi, and M. I. El-Gohary, "Characterization of lung's emphysema distribution: Numerical assessment of disease development," *NISS2010 - 4th Int. Conf. New Trends Inf. Sci. Serv. Sci.*, pp. 464–469, 2010, [Online]. Available: www.emphysemafoundation.org.
- [8] J. Peng *et al.*, "Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging," *Eur. Radiol.*, vol. 30, no. 1, pp. 413–424, 2020, doi: 10.1007/s00330-019-06318-1.
- [9] H. Sugimori, K. Shimizu, H. Makita, M. Suzuki, and S. Konno, "A comparative evaluation of computed tomography images for the classification of spirometric severity of the chronic obstructive pulmonary disease with deep learning," *Diagnostics*, vol. 11, no. 6, 2021, doi: 10.3390/diagnostics11060929.
- [10] C. Hatt, C. Galban, W. Labaki, E. Kazerooni, D. Lynch, and M. Han, "Convolutional neural network based COPD and emphysema classifications are predictive of lung cancer diagnosis," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11040 LNCS, pp. 302–309, doi: 10.1007/978-3-030-00946-5_30.
- [11] H. S. Ryosuke Kobayashi, Toshiyuki Tanaka¹¹ Hidetoshi Nakamura, Toru Shirahata, "Alogirthm of pulmonary Emphysema analysis using Comparing with Expiratory and Inspiratory state of CT images," pp. 3105–3109, 2008, [Online]. Available: www.emphysemafoundation.org.
- [12] L. Peng *et al.*, "Classification and Quantification of Emphysema Using a Multi-Scale Residual Network," *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 6, pp. 2526–2536, 2019, doi: 10.1109/JBHI.2018.2890045.
- [13] C. R. Hatt, A. S. Oh, N. A. Obuchowski, J. P. Charbonnier, D. A. Lynch, and S. M. Humphries, "Comparison of ct lung density measurements between standard full-dose and reduced-dose protocols," *Radiol. Cardiothorac. Imaging*, vol. 3, no. 2, 2021, doi: 10.1148/rct.2021200503.
- [14] S. M. Humphries *et al.*, "Deep learning enables automatic classification of emphysema pattern at CT," *Radiology*, vol. 294, no. 2, pp. 434–444, 2020, doi: 10.1148/radiol.2019191022.
- [15] J. Song *et al.*, "Generative method to discover emphysema subtypes with unsupervised learning using lung macroscopic patterns (LMPS): The MESA COPD study," *Proc. - Int. Symp. Biomed. Imaging*, pp. 375–378, 2017, doi: 10.1109/ISBI.2017.7950541.
- [16] Y. Wu, S. Qi, Y. Sun, S. Xia, Y. Yao, and W. Qian, "A vision transformer for emphysema classification using CT images," *Phys. Med. Biol.*, vol. 66, no. 24, p. 245016, 2021, doi: 10.1088/1361-6560/ac3dc8.
- [17] L. Sørensen, S. B. Shaker, and M. De Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," *IEEE Trans. Med. Imaging*, vol. 29, no. 2, pp. 559–569, 2010, doi: 10.1109/TMI.2009.2038575.
- [18] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017, ChestX-ray8Hospital-ScaleChestCVPR2017_paper.pdf
- [19] NIH News release: NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community
- [20] Original source files and documents: <https://nihcc.app.box.com/v/ChestXRay-NIHCC/folder/36938765345>
- [21] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, and S. Singh, "Deep Transfer Learning Based Classification Model for COVID-19 Disease," *Irbm*, vol. 1, pp. 1–6, 2020, doi: 10.1016/j.irbm.2020.05.003.
- [22] E. Tuba, I. Strumberger, N. Bacanin, and M. Tuba, "Analysis of local binary pattern for emphysema classification in lung CT image," *Proc. 11th Int. Conf. Electron. Comput. Artif. Intell. ECAI 2019*, no. June 2020, 2019, doi: 10.1109/ECAI46879.2019.9042056.
- [23] O. Er and F. Temurtas, "A study on chronic obstructive pulmonary disease diagnosis using multilayer neural networks," *J. Med. Syst.*, vol. 32, no. 5, pp. 429–432, Oct. 2008, doi: 10.1007/s10916-008-9148-6.
- [24] A. Isaac, H. K. Nehemiah, A. Isaac, and A. Kannan, "Computer-Aided Diagnosis system for diagnosis of pulmonary emphysema using bio-inspired algorithms," *Comput. Biol. Med.*, vol. 124, no. March, p. 103940, 2020, doi: 10.1016/j.combiomed.2020.103940.
- [25] C. M. Bhuma, "An Ensemble Approach to Emphysema Classification," *2020 IEEE 15th Int. Conf. Ind. Inf. Syst. ICIIS 2020 - Proc.*, no. 978, pp. 150–155, 2020, doi: 10.1109/ICIIS51140.2020.9342718.
- [26] G. Altan, Y. Kutlu, and N. Allahverdi, "Deep Learning on Computerized Analysis of Chronic Obstructive Pulmonary Disease," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 5, pp. 1344–1350, 2020, doi: 10.1109/JBHI.2019.2931395.
- [27] Y. F. Riti, H. A. Nugroho, S. Wibrama, B. Windarta, and L. Choridah, "Feature extraction for lesion margin characteristic classification from CT Scan lungs image," *Proc. - 2016 1st Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2016*, pp. 54–58, 2016, doi: 10.1109/ICITISEE.2016.7803047.

Detection and Classification of Lung Cancer Using VGG-16

Dr. K. Ramanjaneyulu
Department of Electronics and Communication
Engineering
P.V.P Siddhartha Institute of Technology
Vijayawada, India
kongara.raman@gmail.com,
[\(orcid.org\)](http://orcid.org/0000-0003-0711-0547)

K. Hemanth Kumar
Department of Electronics
and Communication
Engineering
P.V.P Siddhartha Institute of
Technology
Vijayawada, India
hemanthkumar.sv2@gmail.co

K. Snehit,
Velagapudi Ramakrishna
Siddhartha Engineering College
Vijayawada, India
snehith529@gmail.com

G.Jyothirmai
Department of Electronics and Communication Engineering
P.V.P Siddhartha Institute of Technology
Vijayawada, India
jyogarlapati09@gmail.com

K. Venkata Krishna
Department of Electronics and Communication
Engineering
P.V.P Siddhartha Institute of Technology
Vijayawada, India
venkatkrishna021@gmail.com

Abstract— One of the main sources of fatalities in the world is lung cancer. Lung cancer is responsible for 7.6 million deaths worldwide each year, as per the statistics of World Health Organization (WHO). It is only possible to treat lung cancer when it is identified at an early stage. Lung cancer may be diagnosed using a variety of technologies, including isotope, MRI, CT, as well as X-ray. CT scan images are not easy to understand, but using CNN with Image Segmentation is a straightforward approach to detect Lung cancer. CNN (Convolutional Neural Network) is a deep structured technique that has been extensively used to investigate the potential to extract and visualize hidden texture information from image datasets. The objective of the examination is to consequently extricate self-learned components with a start to finish learning CNN and contrast the discoveries with the presentation of standard best in class and customary PC helped indicative frameworks.

Keywords— lung cancer, Image Segmentation, CT Scan, CNN.

I. INTRODUCTION

Cellular breakdown in the lungs is among the riskiest sorts of malignant growth as far as death. Then again, early location of cellular breakdown in the lungs expands the possibilities of endurance. Little cell developments in the lungs called pneumonic knobs can be threatening (destructive) or non - carcinogenic (harmless). The significance of early location of threatening growth knobs is significant for a decent prognosis [1]. Introductory threatening lung knobs impersonate non-dangerous knobs, requiring a differential determination in view of area, underlying contrasts, and clinical biomarkers [2]. The troublesome issue is for deciding the opportunity of threat in early threatening lung nodules [3]. Doctors use a scope of symptomatic ways to deal with identifying dangerous lung knobs early, including clinical conditions, CT ("Computer Tomography") examination assessment (morphological assessment), needle prick biopsy evaluation, and PET sweep examination (metabolic assessments) [4]. Interestingly, medical services suppliers as often as possible utilize intrusive systems, for example, biopsies or medical procedures to recognize threatening or harmless lung knobs.

Obtrusive tasks are risky for a particularly sensitive organ, and they add to patients' anxiety.

CT checking is the best method for researching lung problems[5]. CT checks, then again, convey a higher gamble of misleading positive outcomes along with radiation-related disease. The radiation contact force of low-portion CT is significantly lower than that of standard-portion CT. As indicated by the information, there is no huge contrast in recognition awareness among standard-portion CT and low-portion CT checks. In contrast with chest radiography, disease related mortality were fundamentally diminished in the picked associate treated to low-portion CT examination, as per the NLST ("National Lung Screening Trial") data set. The recognition awareness of lung knobs improves with upgraded anatomic designs (more slender cuts) and better picture enlistment processes. Notwithstanding, the datasets are enormously extended because of this.Based on the cut thickness, up to 500 segments/cuts can be made in a solitary output. A capable radiologist can notice a solitary cut in 2-3.5 minutes. A radiologist is put under a great deal of tension when the person is investigating a CT examination for the presence of a knob. The awareness of knobs to be identified is impacted by various qualities, including their size, position, structure, closeness to different designs, thickness, and edges.

II. PROPOSED SYSTEM

A. Dataset

The Luna16 dataset is used here, which is a registry with a few subdirectories named after the patients' ids. A total subfolder is a 3D picture of the lungs that is isolated into around 180 2D picture cuts in light of their picture number. Sample image of original do com slices is shown in Fig 1. The perception of the dataset is a significant part of the preparation interaction since it helps cognizance. CT check pictures, then again, are challenging to envision on a standard PC or in any window program. Thus, we'll utilize the pydicom library to handle this issue.

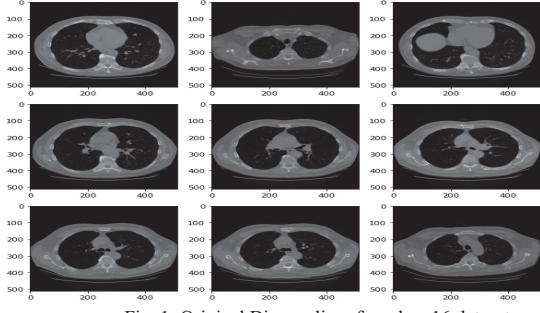


Fig. 1. Original Dicom slices from luna16 dataset

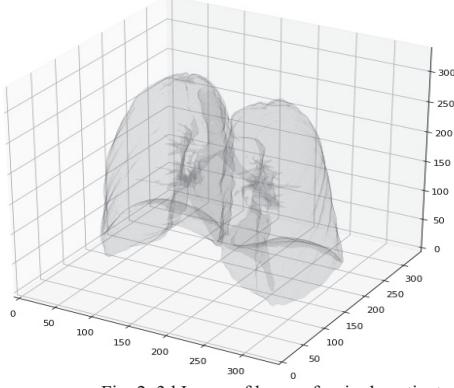


Fig. 2. 3d Image of lungs of a single patient

The Pydicom library gives a picture exhibit as well as metadata data, for example, the patient's ID, name, birth date, picture position, picture number, specialist's name, and specialist's introduction to the world date that is saved in CT checks. A 3d picture is made by consolidating all photographs in a solitary subdirectory and arranging them by picture. The 3D imaging of the lungs gives an exhaustive perspective on cellular breakdown in the lungs cells as well as different diseases. The situation is depicted in Figure 2.

B. Segmentation

Segmenting a picture by using the watershed technique may be a classic use of this algorithm. A watershed is a transition that may be seen in a grayscale picture. Watershed algorithms will make our task easier and effective. The watershed change regards the picture as a geological guide, with the brilliance of every pixel indicating its rise, and tracks down the lines that go over the highest points of edges. After assigning a set of markers to a watershed line, the algorithm begins flooding the basins associated with that line. Markers are frequently used as the image's local minima, which are subsequently used to flood basins in the image.

In the first place, utilizing twofold enlargements, inward and outer markers are recovered from CT examine pictures and joined with a totally dark picture using watershed draws near. It likewise diminishes picture clamor and gives a watershed marker to lungs and malignant growth cells. The watershed marker diminishes outer clamor and applies a twofold cover to the picture, as found in fig 3; dark lung pixels indicate malignant growth cells.

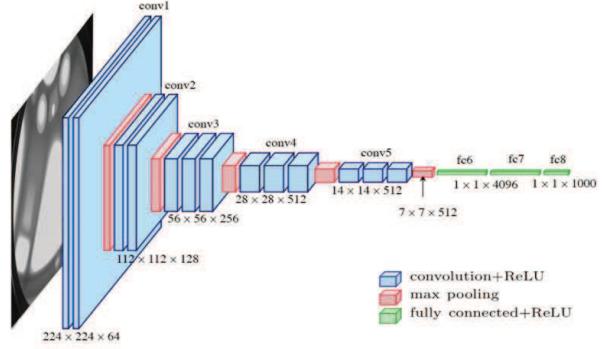


Fig 3 VGG-16 Architecture

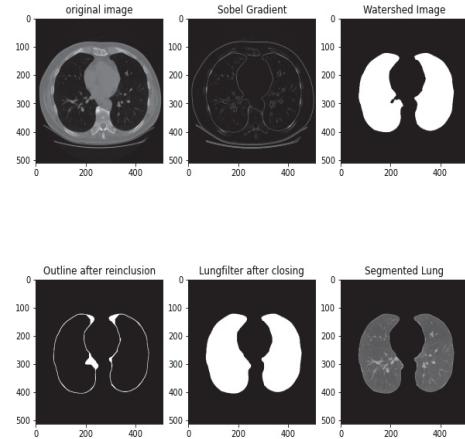


Fig. 4. Image segmentation process visualization

The Sobel filter is combined with watershed algorithms to improve segmentation. The outer layer of the lungs is removed. We utilize the interior marker and the Outline that was simply built to produce the lung channel utilizing bitwise or NumPy activities in the wake of eliminating the external layer. The heart is likewise removed from CT scan images. The next stage is to use morphological operations and morphological gradients to block off the lung filter. It results in better-segmented lungs than the previous method. Figure 4 depicts the procedure.

Pre-processing of the input image including segmentation of regions such as the lungs and cancer areas will be employed to improve the model's accuracy for processing and result finding. For segmentation, additional networks such as the U-Net as well as Nested U-Net will be used.

C. Proposed Model

The CNN that will be utilized in the proposed technique is prepared utilizing lung division from CT check pictures. We start by preprocessing luna16's dataset. In the wake of preprocessing, the following stage is to utilize a watershed calculation to segment the lungs. Utilizing the semantic division approach, the Watershed calculation stresses the lung segment and makes parallel veils for lungs. The model strategy was to utilize move learning on VGG-16 for certain changes in the last three completely associated layers. This

model gives considerable outcomes in object characterization. Fig 5 shows the model's synopsis.

Model: "vgg16"		
Layer (type)	output Shape	Param #
input_3 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
fc3 (Dense)	(None, 128)	524416
fc4 (Dense)	(None, 128)	16512
output (Dense)	(None, 1)	129
<hr/>		
Total params:	134,801,601	
Trainable params:	17,322,369	
Non-trainable params:	117,479,232	

Fig. 5. Model summary of Transfer learning model-VGG16

VGG Net has been educated to remove highlights (include extractor) that can distinguish things and is presently being utilized to arrange objects that aren't apparent. VGG 16 and 19, which contain 16 and 19 weight layers, separately, are utilized for object recognizable proof. VGG Net plays out a pile of convolutional layers utilizing 224*224 RGB pictures as contribution with a decent 33% channel size and a step of 1. For down sampling the info portrayal, convolutional layers are joined with five max-pooling channels. The pile of convolutional layers is trailed by three completely connected layers, each with 4096, 4096, and 1000 channels. The last layer is Soft-max. The organization construction of VGG is displayed in the outline underneath. Nonetheless, in this methodology, photos having the state of (512x512) are considered. Accordingly, another model is made using the vgg16-net design. What's more, I utilized a strong Adam streamlining agent to develop the model; the learning rate is 0.0001, the entropy is paired cross-entropy, and the precision measurements are

generally parallel cross-entropy. A model outline, convolution layers, max-pooling layers, and params are displayed in Figure 6.

Model: "vgg16"		
Layer (type)	output Shape	Param #
input_3 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
predictions (Dense)	(None, 1000)	4097000
<hr/>		
Total params:	138,357,544	
Trainable params:	138,357,544	
Non-trainable params:	0	

Fig 6. Model summary

III. RESULTS

Fig 7 illustrates Validation and training accuracy and Fig 8 depicts validation and training loss.

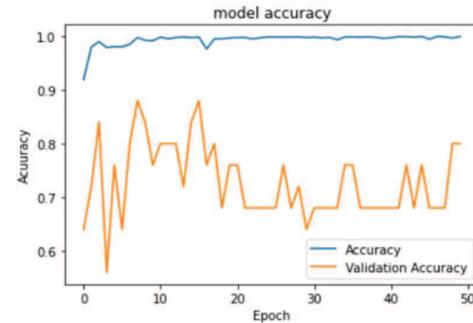


Fig. 7. Validation Accuracy and Training of model VGG_16

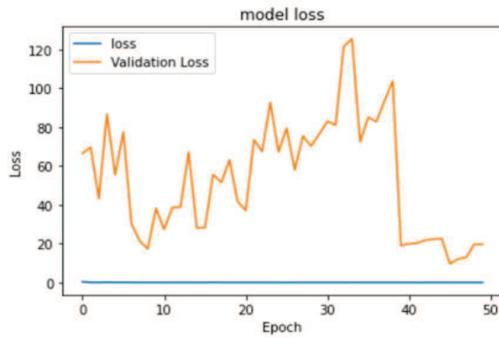


Fig 8. Validation loss and Training of model VGG_16

As the callbacks capacity will store hands down the best precision model, unquestionably the best val. acc. related data in all models will be saved. Table 1 analyzes model preparation precision and misfortune to display approval exactness and misfortune.

TABLE I. ACCURACY AND LOSS COMPARISON OF DIFFERENT MODELS

Index	Model	Train Acc.	Train loss	Val. Acc.	Val. loss
1.	VGG_16	99.84%	0.0046	88.00%	28.2614

A. Conclusion

The model built for recognizing cellular breakdown in the lungs utilizing VGG 16 is performing admirably while preparing the model and acquiring adequate degrees of test precision and test misfortune, as indicated by the outcomes introduced in Adobe diagrams and the correlation table. The model can be deployed to predict possible malignant tumors which are in early stages.

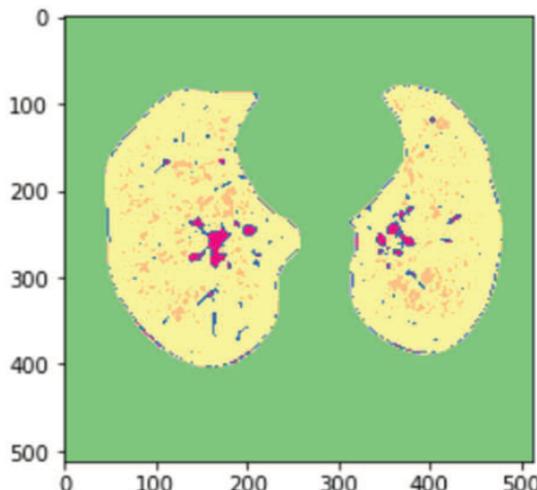


Fig 9. Image of Lung cancer plotted with Matplotlib

REFERENCES

- [1] Bjerager M., Palshof T., Dahl R., Vedsted P., Olesen F. Delay in the diagnosis of lung cancer in general practice. *Br. J. Gen. Pract.* 2006;56:863–868.
- [2] Nair M., Sandhu S.S., Sharma A.K. Cancer molecular markers: A guide to cancer detection and management. *Semin. Cancer Biol.* 2018;52:39–55. DOI: 10.1016/j.semcaner.2018.02.002.
- [3] Silvestri G.A., Tanner N.T., Kearney P., Vachani A., Massion P.P., Porter A., Springmeyer S.C., Fang K.C., Midthun D., Mazzone P.J. Assessment of plasma proteomics biomarker's ability to distinguish benign from malignant lung nodules: Results of the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) trial. *Chest.* 2018;154:491–500. DOI: 10.1016/j.chest.2018.02.012.
- [4] Shi Z., Zhao J., Han X., Pei B., Ji G., Qiang Y. A new method of detecting pulmonary nodules with PET/CT based on an improved watershed algorithm. *PLoS ONE.* 2015;10:e0123694.
- [5] Lee K.S., Mayo J.R., Mehta A.C., Powell C.A., Rubin G.D., Prokop C.M.S., Travis W.D. Incidental Pulmonary Nodules Detected on CT Images: Fleischner 2017. *Radiology.* 2017;284:228–243.
- [6] Diederich S., Heindel W., Beyer F., Ludwig K., Wormanns D. Detection of pulmonary nodules at multirow-detector CT: Effectiveness of double reading to improve sensitivity at standard-dose and low-dose chest CT. *Eur. Radiol.* 2004;15:14–22.
- [7] Demir Ö., Çamurcu A.Y. Computer-aided detection of lung nodules using outer surface features. *Bio-Med. Mater. Eng.* 2015;26:S1213–S1222. DOI: 10.3233/BME-151418.
- [8] Bogoni L., Ko J.P., Alpert J., Anand V., Fantauzzi J., Florin C.H., Koo C.W., Mason D., Rom W., Shiu M., et al. Impact of a computer-aided detection (CAD) system integrated into a picture archiving and communication system (PACS) on reader sensitivity and efficiency for the detection of lung nodules in thoracic CT exams. *J. Digit. Imaging.* 2012;25:771–781. DOI: 10.1007/s10278-012-9496-0.
- [9] Al Mohammad B., Brennan P.C., Mello-Thoms C. A review of lung cancer screening and the role of computer-aided detection. *Clin. Radiol.* 2017;72:433–442. DOI: 10.1016/j.crad.2017.01.002.
- [10] Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies. Nasraullah Nasrullah, Jun Sang, Mohammad S. Alam, Muhammad Mateen, Bin Cai and Haibo Hu.
- [11] Setio A.A.A., Traverso A., de BelT., Berens M.S.N., van den Bogaard C., Cerello P., Chen H., Dou Q., Fantacci M.E., Geurts B., et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Med. Image Anal.* 2017;42:1–13. DOI: 10.1016/j.media.2017.06.015.
- [12] Zhu W., Liu C., Fan W., Xie X. DeepLung: Deep 3D dual-path nets for automated pulmonary nodule detection and classification; Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV); Lake Tahoe, NV, USA. 12–15 March 2018; pp. 673–681.
- [13] Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017;39:1137–1149. DOI: 10.1109/TPAMI.2016.2577031.
- [14] Jiang H., Ma H., Qian W., Gao M., Li Y. An automatic detection system of lung nodules based on a multigroup patch-based deep learning network. *IEEE J. Biomed. Heal. Inform.* 2018;22:1227–1237. DOI: 10.1109/JBHI.2017.2725903.
- [15] Masood A., Sheng B., Li P., Hou X., Wei X., Qin J., Feng D. Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images. *J. Biomed. Inform.* 2018;79:117–128. DOI: 10.1016/j.jbi.2018.01.005.

Credit Card Fraud Detection with Unbalanced Real and Synthetic dataset using Machine Learning models

Janet B
Department of Computer Application
NIT Tiruchirappalli
Tiruchirappalli, India
janet@nitt.edu

Joshua Arul Kumar R
Department of Mechatronics
Engineering
Rane Polytechnic Technical Campus
Tiruchirappalli, India
r.joshuaarulkumar@ranepolytechnic.ed
u.in

Didugu Phani Sai Ganesh
Computer Applications
National Institute of Technology
Tiruchirappalli
Tamil Nadu
ganeshdidugu6@gmail.com

Abstract—As credit card becomes one of the most trusted and popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. It is of utmost importance for financial institutions like banks and credit card companies to find fraudulent credit card transactions in real time so as to withhold any suspicious transaction till they get further confirmation from the customers and ensure customers aren't paying for anything they haven't bought. But the challenge in making such a model that classifies transactions into legitimate and fraudulent transactions is that the legitimate transactions occur much more frequently than the fraud transactions. In this paper we will analyse how the traditional machine learning algorithms handle highly imbalanced data, biased against fraud transactions, and compare it with algorithms that are designed specifically to deal with highly imbalanced data. For fraud detection in mobile money transaction, a few supervised machine learning models (Logistic Regression, Random Forest, Decision Tree, SVM) are tested and compared. All classification models are tested using a synthetic dataset generated by a simulator based on real transaction of a company.

Keywords—*fraud, random forest, precision, recall, smote, fraud detection, supervised learning, logistic regression, decision tree, SVM*

I. INTRODUCTION

In 2018, unauthorized financial fraud losses across payment cards and remote banking added up to £844.8 million in UK. On the bright side financial institutions have prevented £1.66 billion from unauthorized fraud in the same year. £2 out of every £3 of fraud were prevented proactively [1]. Because of continuous growth in technology, financial transactions by mobile applications have increased rapidly. It is an easy and convenient way for trade [2] between customer and merchants, bank transfer, and transferred between users. Money transaction through UPI (Unified Payment Interface) has increased rapidly in recent years in India. Volume of UPI transactions in India in September was 3.65 billion worth 6.54 trillion rupees.

Despite convenience and increasing popularity, fraud in mobile transactions is increasing day by day. In the first quarter (Jan – Apr) 2021 compared to last quarter of 2020, fraud attempt has increased by 89% in India, and by 149% globally. Losses due to online payment fraud is doubling every year. In 2015, fraudsters caused PayPal a \$1 million fraud scheme loss. Therefore, an effective detection system will help companies in their fight with fraudsters [3].

II. DATASET

In this paper, we will be working with two datasets. The first is the Credit Card Fraud Detection dataset downloaded from Kaggle. The dataset contains transactions made by credit cards in September 2013 by European cardholders. This dataset presents transactions that occurred in two days, where we have 492 frauds out of 284,807 transactions [4].

It contains only numerical input variables, most of which are the result of a Principle component analysis (PCA) transformation done to maintain the privacy of the cardholders. Columns Time, Amount and Class are not the only 3 columns that are not transformed with PCA.

Time column gives the value of time passed (in seconds) since we started collecting the transaction data. Amount column gives the amount that is being transacted. Class column is our target variable. If class is '1' it means that the particular transaction is a fraud transaction and if it is '0', then the particular transaction is a legitimate transaction. They are 284315, 492 respectively.

The feature 'Time' contains the seconds elapsed between each transaction and the first transaction in the dataset. The feature 'Amount' is the transaction amount. Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions.

First, we divide the dataset into two parts: training set and test set. Training set is used to train the model, after training the model, we use the test set to predict outputs. Then we compare the predicted output with actual output to evaluate the model. In this study, we using Logistic Regression, Random Forest and Support Vector Machine Classifier algorithms to predict the output.

Using correlation matrix, we can find the significance of a feature. If any features do not have any correlation with the dependent variable, we can drop the feature. Here in our dataset isFlaggedFraud does not have any correlation with isFraud, so we can drop isFlaggedFraud. oldbalanceOrg and type are mostly correlated with isFraud. There are also some features, which are negatively correlated with isFraud.

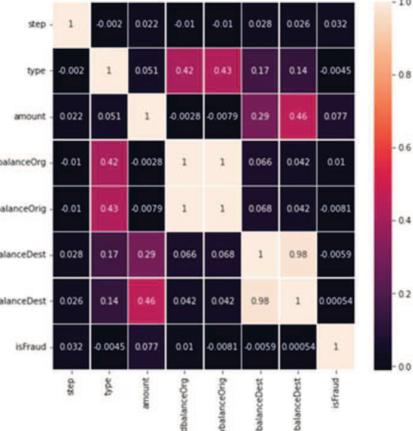


Fig. 1. correlation matrix

The second dataset is a synthetic dataset [4] generated by a simulator based on real transactions of a company. The dataset contains 6362620 records of transaction and 11 features. The features are-

step: 1 step is equal to 1 hour of time. The dataset has 744 step means total 30 days of simulation.

type: Type of transaction, type has 5 categorical value, DEBIT, PAYMENT and TRANSFER, CASH-IN, CASH-OUT.

amount: Amount of the transaction in local currency.

nameOrig: Customer who started the transaction.

oldbalanceOrg: Initial balance before the transaction.

newbalanceOrg: New balance after the transaction.

nameDest: Customer who is recipient of the transaction.

oldbalanceDest: Initial balance of the recipient before the transaction.

newbalanceDest: New balance of the recipient after the transaction.

isFraud: This is the transaction made by the fraudulent agent inside the simulation. It has two values, 0 means not a fraud transaction, and 1 means this is a fraud transaction.

isFlaggedFraud: If the amount of transaction is greater than 200,000, then it will raise a flag.

Here, type is an object variable, so it is converted into int64 by assigning each class an integer value. CASH-IN means cash inflow in the customer's account; CASH-OUT is the opposite of CASH-IN, DEBIT means sending money from mobile to a bank account; PAYMENT means that customer is paying for goods or services to merchants; TRANSFER means transaction between users.

Out of 11 features only 8 features are important, so the features which are not important are dropped: nameOrig and nameDest are all unique value, and isFlaggedFraud is biased so it does not have any relation with isFraud. isFraud is the output feature

III. EVALUATION METRIC

Given the high-class imbalance ratio, accuracy will not give us the best picture of how well our model is performing because no model (predicting every transaction is false) is accurate. To understand this, let's take a look at the confusion matrix of the model as shown in figure 2.

	Predicted_Fraud	Predicted_Not_Fraud
Actual_Fraud	0	98
Actual_Not_Fraud	0	56864

Fig. 2: Confusion matrix of the model.

Even when we predict no transaction is fraud our model is doing great if we consider accuracy which clearly is not desirable. To understand which prediction will work in our case, we can refer to confusion matrix of a typical binary classification that is shown in figure 3.

		Predicted condition	
		Total population = P + N	
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

Fig. 3. Confusion matrix of a binary classification [12]

Accuracy measures the ratio of number of correctly predicted datapoints to total number of datapoints in the dataset. If TN dominates the whole matrix, we can just ignore the rest of the cases if we are looking only at accuracy. But in reality, especially while detecting fraud we need to minimize both FN and FP values while increasing TP. FN indicates the fraud transaction that our model missed and FP indicates the legitimate transactions our model flagged as a fraud transaction.

A. Precision Recall Tradeoff

Precision is defined as the ratio of TP to PP. It gives us the proportion actual fraud transactions among the transactions that are flagged as fraud by our model. Recall is defined as the ratio of TP to P. It gives us the proportion of transaction of that our model flagged fraud of the total number of actual fraud transactions.

It is the situation that the expense of mislabeling a fraud exchange as a genuine exchange is frequently a lot higher than the expense of the converse mistake.

Ideally, we would want a high recall and a high precision. But reality is never that easy. If we try to make sure that our model has a high recall i.e. no fraud transaction is getting through, then it would inevitably lead to low precision, i.e. normal legitimate transactions get suspended because our model flagged them as fraud transaction. The inverse is also true. If we try to make a model that caters much more towards customer experience and hence has a high precision, i.e. no flagged transaction is a legitimate transaction, this would result in fraud transactions going undetected.

The industry standard is to have a recall of 1 and an acceptable precision. Recall of 1 says that no fraud transaction goes unflagged by our model. To achieve this, we have to compromise on precision i.e. some legitimate transactions might also get flagged as fraud transactions. Intuitively, this is

understandable because letting through a fraud transaction is much worse than the inconvenience caused to a customer when their normal transaction is suspended because it is flagged as a fraud transaction by our model. Of course, no one wants very high precision because it will degrade the customer experience. So it's up to the data scientist to prioritize precision or recall.

In our paper, we will follow the industry standard trying to maximize recall while keeping precision in acceptable limits. Sometimes we will also look at another metric called f1 score which is double of harmonic mean of precision and recall. If f1 score is too low, it indicates one or both of precision and recall is too low.

IV. PRELIMINARY DATA ANALYSIS

There are 28 anonymized variables obtained using PCA, named V1 to V28. The variable time gives the time in seconds since the data collection has begun. The amount gives the amount in £ that are being transacted.

The statistics of the variable time are shown in figure 4. It ranges from 1 to 1,72,792. We transformed the variable to show the hour at which a particular transaction happened. Doing that gives us some meaningful insights into when fraud happens.

```
count      284807.000000
mean       94813.859575
std        47488.145955
min        0.000000
25%       54201.500000
50%       84692.000000
75%       139320.500000
max       172792.000000
Name: Time, dtype: float64
```

Fig. 4: Statistics of Time Variable

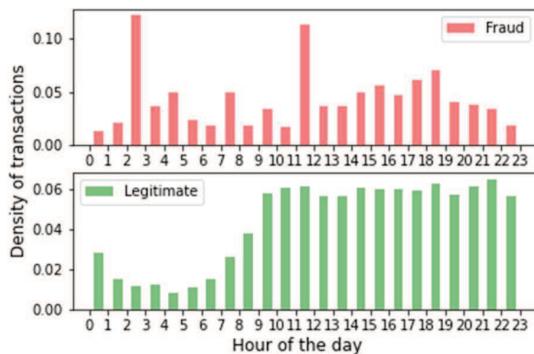


Fig. 5. Distribution of density of transactions throughout the day.

We can see from figure 5 that fraud transactions peak in the middle of the night between 2:00 PM and 3:00 PM and 11:00 AM to 12:00 PM. Such type of information will be useful while training decision trees. Decision trees can pick up such pattern. Also note that the normal transaction density remains approximately constant once the day begins i.e. from 09:00 AM to 23:00 AM.

Figure 6 shows the statistical properties of data in Amount columns. As you can see the transaction amount ranged from £0 to £25,691. Only 25% of transactions have an amount more than £77.165. Half of all transactions [6] are below £22 only. Only 1% of all transactions have a transaction amount greater than £1017.97.

```
count      284807.000000
mean       88.349619
std        250.120109
min        0.000000
25%       5.600000
50%       22.000000
75%       77.165000
max       25691.160000
Name: Amount, dtype: float64
```

Fig. 6. Statistical properties of data 'Amount'

Fortunately, our dataset doesn't have any null values or missing values. This is because our data has been anonymized using PCA which transform null values too.

V. RANDOM UNDER SAMPLING

Random under sampling is a statistical technique commonly used when classes are highly imbalanced. In this technique, we keep randomly eliminating the majority class datapoints till the classes are balanced.

We have created a subsample from our training data set using random under sampling technique. This subsample is then split into training subsample and testing subsample. We trained different algorithms on the training subsample and validated them on testing subsample along with testing data which is highly imbalanced. Validating our models on both testing subsample and testing data gave us much clearer picture of which algorithms work better under random under sampling.

The machine learning algorithms that we have implemented are Logistic Regression [7], Support Vector Classification [8], Decision Tree Classification [9], Random Forest Classification [10].

The performances of all these algorithms after random under sampling the data are summarized in table 1.

TABLE 1 PERFORMANCE OF MODELS TRAINED ON RANDOMLY UNDER SAMPLED DATA

Model	Recall	Precision	f1 score
Logistic Regression	0.9898	0.0028	0.0056
Support Vector Classifier	1.0000	0.0017	0.0034
Decision Tree	0.1553	0.8367	0.0856
Random Forests	0.8367	0.0856	0.1553

The table 1 lists the best performance achieved after tuning their hyperparameters. While tuning the hyperparameters, we tried to maximize the recall score. This can be reflected in case of support vector classifiers where it achieved a recall of 1.000

at the cost of precision. This is what we mean by precision recall trade-off.

It is observed that either models tend to have good recall score or good precision score. What we need is somewhere in between but a little bit towards having a better recall. Random under sampling doesn't necessarily produce the output that we want which is a recall score of 1 and an acceptable precision score.

Logistic regression proved to be very useful in detecting fraud. But its disadvantage is that it also gives a lot of false positives. A better model than this will be Random Forest even though their recall is lower than logistic regression, its precision is much better than the precision of logistic regression.

VI. RANDOM OVER SAMPLING

Random over sampling [11] is another statistical technique in which we generate synthetic samples [12] of minority classes to balance the class imbalance.

Under sampling [13] the majority classes is found to be effective in tuning the models sensitivity towards non-majority classes. [14] shows a mix of over-sampling the minority classes and under-sampling the majority classes will produce much better performance than only using under-sampling or only using over-sampling. It also proposes an over-sampling technique called Synthetic Minority Over-sampling Technique (SMOTE).

A. SMOTE

SMOTE is an over-sampling technique. What SMOTE does is that it takes the minority class from the data set and then oversamples them by synthesizing new data points rather than just repeating the same datapoints from the majority class.

[15] first used a similar technique to successfully create extra training data for handwritten character recognition. In their implementation, rotation and skew operations were natural ways to perturb the training data and still be accurate because handwritten digits rotated and skewed still represent the same digit to some degree.

The minority class is over-sampled by taking each minority class sample and introducing synthetic examples along the line segments joining any/all of the k minority class nearest neighbours. Depending upon the amount of over-sampling required, neighbours from the k nearest neighbours are randomly chosen.

For our practical implementation purposes, we are using an established and well reputed python package called "imbalanced-learn" [16] that is specifically built to deal with highly imbalanced class data [17] [18] [19] like ours.

B. Logistic Regression

We have already seen the performance of logistic regression on randomly under sampled data. It is concluded that logistic regression tends to give a high recall, low precision model.

TABLE 2 PERFORMANCE OF LOGISTIC REGRESSION ON VARIOUS SAMPLING METHODS

Metric	Original Data	Random Under sampled data	Over sampled data using SMOTE
Recall	0.6327	0.9898	0.8673
Precision	0.9118	0.0028	0.5903
F1 score	0.7470	0.0056	0.7025

From table 2 it is clear that logistic regression doesn't work very well with random under sampling method. Model will have high variance and low precision. We lose confidence in the prediction. It is also observed that SMOTE sampling didn't produce the desired result. Over sampling increases recall at the cost of precision which is to be expected. But the improvement in recall doesn't justify the precision trade off and the computation cost involved with SMOTE oversampling. This is also reflected in f1 score.

C. Support Vector Classification

SVC is not suitable for classification of large data set, because the training complexity of SVM [20] is quadratic time complexity with respect to the number of datapoints in the dataset. Our training set has a total of 2,27,845 rows. The accuracy SVC produces doesn't justify the computation cost. Hence, for the purpose of our research, we are not using SVC on the entire training data or on SMOTE over sampled data which is even larger than our training data to begin with.

There are several research works regarding computationally efficient SVC algorithms that are shown to give acceptable performance, but the application of such algorithms is out of the scope of this project. A good review of such algorithms can be found at [21].

D. Decision Tree Classifiers

It can be seen from the results of Table 3, that the decision trees are highly sensitive to sampling strategies in general. Both random under sampling and over sampling hurts the performance of the model, in general both reduce precision and recall scores and thus the overall f1 score. Random under sampling gives much worse results than over sampling techniques. This is due to the information loss that happens in the process of random under sampling.

TABLE 3: PERFORMANCE OF DECISION TREE ON VARIOUS SAMPLING METHODS

Metric	Original Data	Random Under sampled data	Over sampled data using SMOTE
Recall	0.7755	0.8367	0.8469
Precision	0.8000	0.0856	0.3790
F1 score	0.7876	0.1553	0.5237

E. Random Forest Classifiers

Random forest classifiers [22] outperformed all other models in sampling techniques that we performed. Random under sampling technique gave the best recall score but at a huge cost of precision. Oversampling with SMOTE gives a reasonable recall score at much better precision score. Even

though we prefer a model with high recall score, we should keep in mind that the precision should be acceptable which it is not in random under sampling data.

TABLE 4: PERFORMANCE OF RANDOM FOREST ON VARIOUS SAMPLING METHODS

Metric	Original Data	Random Under sampled data	Over sampled data using SMOTE
Recall	0.7245	0.8367	0.8061
Precision	0.9467	0.0856	0.9294
F1 score	0.8208	0.1553	0.8634

In the second synthetic dataset, all the algorithms gave high accuracy, because 99% of the sample data is not fraud (the data is synthetic), the prediction in case of not fraud will be true maximum of the time, hence the accuracy will be very high. Logistic Regression takes less time than Random Forest algorithm, but the F-1 score for Random Forest is higher than the Logistic regression. Support Vector classifier takes lot of time to fit the train data and the f-score is lower than the Random Forest algorithm as shown in table 5.

TABLE 5: PERFORMANCE OF SYNTHETIC DATA ON VARIOUS MACHINE LEARNING MODELS

Metric	Logistic regression	Random forest	Support vector machine
Accuracy	99.82	99.97	99.8
F1 score	39.52	87.6	51.8

The synthetic dataset is highly imbalanced. It has 6 million data in majority class and only 8 thousand data in minority class. Because of this imbalance in dataset accuracy of our models is not very good. Here we are using Random Under Sampler to reduce the number of records in majority class. We are taking only 821300 records out of 6 million. we are also using Random Over Sampler to increase the number of records in minority class. We are increasing the number of records 8213 to 635440 in minority class, which is 10% of majority class.

TABLE 6: PERFORMANCE OF SYNTHETIC DATA ON RANDOM OVER/UNDER SAMPLING METHODS

Metric	Random Under sampled data		Random Over sampled data	
	Logistic regression	Random forest	Logistic regression	Random forest
Accuracy	99.12	99.66	99.97	99.99
F1 score	56.74	92.72	76.67	75.97

Using RandomUnderSampler, Random Forest is giving F1-score of 92.72% and Logistic Regression is giving 56.74%,

which is better than previous Logistic Regression Model. And using RandomOverSampler, Random Forest is giving F1-score of 99.97% and 0 false negative.

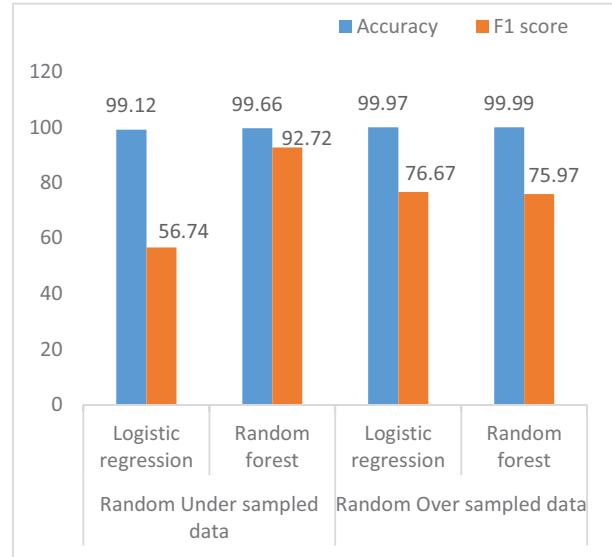


Fig. 6. Performance of synthetic data on random over/under sampling

So, Random Forest algorithm using Random Over Sampler is our best model for transaction fraud detection as shown in figure 6.

CONCLUSION

In this paper, we have analyzed the highly unbalanced real and synthetic dataset and discussed the various metrics that we can use to evaluate the models. It is concluded to use recall score while keeping an eye on precision score. Random under sampling technique is used to balance the classes and several models have been built on that sampled data. It is concluded that Random Forest Classifiers gives the best performance on random under sample data. SMOTE technique is used to perform over sampling of minority classes. It is concluded that Random Forest classifiers applied on SMOTE oversampled data gives the best performance results. It is to be expected because random forest classifiers are very powerful and generalized models that are well equipped to deal with imbalanced classes.

Further studies can be conducted to evaluate how multilayer perceptron and deep learning models will perform on such highly imbalanced real and synthetic datasets.

REFERENCES

- [1] "FRAUD THE FACTS 2019 - The definitive overview of payment industry fraud". UK Finance. <https://www.ukfinance.org.uk/system/files/Fraud%20The%20Facts%202019%20-%20FINAL%20ONLINE.pdf> accessed on 29.3.22
- [2] Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson and Gianluca Bontempi. Calibrating Probability with Undersampling for Unbalanced Classification. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015
- [3] Dal Pozzolo, Andrea; Caelen, Olivier; Le Borgne, Yann-Ael; Waterschoot, Serge; Bontempi, Gianluca. Learned lessons in credit card fraud detection from a practitioner perspective, Expert systems with applications, 41,10,4915-4928,2014, Pergamon
- [4] <https://www.kaggle.com/mlg-ulb/creditcardfraud>
- [5] T. M. Khoshgoftaar, M. Golawala and J. V. Hulse, "An Empirical Study of Learning from Imbalanced Data Using Random Forest," 19th

- IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007), 2007, pp. 310-317, doi: 10.1109/ICTAI.2007.46.
- [6] E. Osuna, R. Freund and F. Girosi, "An improved training algorithm for support vector machines," Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop, 1997, pp. 276-285, doi: 10.1109/NNSP.1997.622408.
 - [7] Ganganwar, Vaishali. "An overview of classification algorithms for imbalanced datasets." International Journal of Emerging Technology and Advanced Engineering 2.4 (2012): 42-47.
 - [8] Yann-Aël Le Borgne, Gianluca Bontempi Machine Learning for Credit Card Fraud Detection - Practical Handbook
 - [9] Dal Pozzolo, Andrea; Boracchi, Giacomo; Caelen, Olivier; Alippi, Cesare; Bontempi, Gianluca. Credit card fraud detection: a realistic modeling and a novel learning strategy, IEEE transactions on neural networks and learning systems,29,8,3784-3797,2018,IEEE
 - [10] Carcillo, Fabrizio; Dal Pozzolo, Andrea; Le Borgne, Yann-Aël; Caelen, Olivier; Mazzer, Yannis; Bontempi, Gianluca. Scarff: a scalable framework for streaming credit card fraud detection with Spark, Information fusion,41, 182-194,2018,Elsevier
 - [11] Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson and Gianluca Bontempi. Calibrating Probability with Undersampling for Unbalanced Classification. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015
 - [12] Dal Pozzolo, Andrea Adaptive Machine learning for credit card fraud detection ULB MLG PhD thesis (supervised by G. Bontempi)
 - [13] Carcillo, Fabrizio; Le Borgne, Yann-Aël; Caelen, Olivier; Bontempi, Gianluca. Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization, International Journal of Data Science and Analytics, 5,4,285-300,2018,Springer International Publishing
 - [14] Fabrizio Carcillo, Yann-Aël Le Borgne, Olivier Caelen, Frederic Oblé, Gianluca Bontempi Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection Information Sciences, 2019
 - [15] N. V. Chawla, K. W. Bowyer, L. O.Hall, W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of artificial intelligence research, 321-357, 2002.
 - [16] Fernández, Alberto, et al. "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary." Journal of artificial intelligence research 61 (2018): 863-905.
 - [17] Ha, T. M., & Bunke, H. (1997). Off-line, Handwritten Numeral Recognition by Perturbation Method. Pattern Analysis and Machine Intelligence, 19/5, 535–539.
 - [18] Bertrand Lebichot, Yann-Aël Le Borgne, Liyun He, Frederic Oblé, Gianluca Bontempi Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection, INNSBDDL 2019: Recent Advances in Big Data and Deep Learning, pp 78-88, 2019
 - [19] Lemaître, Guillaume, Fernando Nogueira, and Christos K. Aridas. "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning." The Journal of Machine Learning Research 18.1 (2017): 559-563.
 - [20] Menon, Aditya Krishna. "Large-Scale Support Vector Machines: Algorithms and Theory." (2009).
 - [21] J. Cervantes, X. Li and W. Yu, "SVM Classification for Large Data Sets by Considering Models of Classes Distribution," 2007 Sixth Mexican International Conference on Artificial Intelligence, Special Session (MICAI), 2007, pp. 51-60, doi: 10.1109/MICAI.2007.27.
 - [22] Dal Pozzolo, Andrea; Caelen, Olivier; Le Borgne, Yann-Ael; Waterschoot, Serge; Bontempi, Gianluca. Learned lessons in credit card fraud detection from a practitioner perspective, Expert systems with applications,41,10,4915-4928,2014, Pergamon

An Anomaly Based Network Intrusion Detection System using LSTM and GRU

Rachana Koniki

*Department of Computer Science
and Engineering
VR Siddhartha Engineering
College, Kanuru,
India
konikirachana@gmail.com*

Mounika Durga Ampapurapu

*Department of Computer Science
and Engineering
VR Siddhartha Engineering
College, Kanuru,
India
mounika9301@gmail.com*

Praveen Kumar Kollu

*Department of Computer Science
and Engineering
VR Siddhartha Engineering
College, Kanuru,
India
praveenman@gmail.com*

Abstract-- Today, billions of devices are connected to the internet and the count keeps on increasing. Most of these networked devices are vulnerable to security attacks. Various types of active and passive attacks will cause severe damage to the privacy and security of millions of users. Even though we have firewalls for security but they limit the access between the networks to prevent intrusion, but they do not alert when there is an attack inside a network. The proposed method uses NSL-KDD dataset, a benchmark dataset from the Canadian Institute for Cybersecurity. We created a network intrusion detector using this dataset, which is a prediction model capable of distinguishing between "bad" connections, which are subsequently categorised into the classifications DoS, Probe, and R2L, and "good" normal connections. This study offers a Deep Learning-based Network Intrusion System. After training the model we achieve a good accuracy and precision. The highest accuracy of about 96% for classifying the Probe Attack. We got 92% accuracy for DOS attack and 88% for R2L.

Keywords: Intrusion Detection, NSL-KDD, Cyber-attacks, LSTM, GRU.

I. INTRODUCTION

Developing a perfect computer system which is absolutely secure is not possible. Most of the existing systems have vulnerabilities and they can be exploited by intruders or even insiders. We can't always know what kinds of attacks are possible and write the rules to detect them every single time. There is always a possibility of discovering a new kind of threat so firewalls and other similar software are not suitable all the times. So, we need a software which can detect the possible malicious intrusions. Early detection of intrusion can save lots of efforts and funds. Considering the possible capabilities of an anomaly based intrusion detection system(network), this approach is currently a significant focus of intrusion detection research and development. A-NIDS-capable systems are now more widely available, and several plans are being investigated. However, the concept is far from mature, and significant concerns must be resolved before A-NIDS platforms can be used on a large scale. Another critical barrier with anomaly based network intrusion detection is the scarcity of labelled data for model training and validation.

Normal conduct labels are frequently accessible, whereas intrusion labels are not. As a result, in this scenario, semi-supervised and unsupervised anomaly detections are selected.

A. DEEP LEARNING

Deep Learning [1] is made up of number of processing layers that work together in order to learn data representations at various degrees of abstraction. Using Deep learning we can identify detailed structure in big data sets by using the technique known to be backpropagation to figure out how to adjust the internal parameters of a machine that are used to calculate the representation in each layer from the preceding layer's representation

1.1 RECURRENT NEURAL NETWORK (RNN)

RNN is one form of neural network that accepts sequential input, such as voice or language. It keeps track of the preceding items in the sequence in recurrent neural networks [2]. They're effective, but they're difficult to train since the back propagated gradients might expand or reduce at each step. They'll either explode or vanish. The primary purpose of a RNNs is to acquire long term dependencies, however RNNs are challenging to master in terms of storing data for lengthy periods of time.

1.2 LONG SHORT TERM MEMORY (LSTM)

LSTM is the short form of Long Short-Term Memory. It consists of special hidden units and remembers the inputs for a long time [3]. The LSTM networks are more efficient than the current neural networks when they have many layers for each step. A cell, an output gate, a forget gate and an input gate make up each LSTM unit. LSTM networks are best suited for classifications and predictions. Initial LSTMs are developed to solve vanishing gradient problem. LSTMs are widely used. Using LSTM, we can add or remove information to the cell states and they are regulated carefully by gates.

1.3 GATED RECURRENT UNIT (GRU)

A GRU is an enhanced form of a Recurrent Neural Network designed to tackle the vanishing gradient issue [4]. GRU is made up of two gates: the update gate and the reset gate. The key benefit of GRU is that they are trained to preserve information for long periods of time without discarding information that is unrelated to predictions. The update gate specifies how much previous data must be saved for the future. The reset gate determines how much past information to discard.

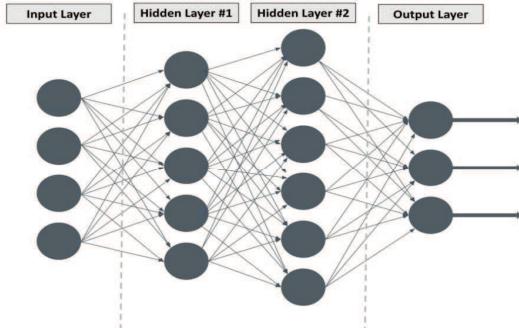


Figure 1. Representation Deep learning model

B. Types of Attacks

a. Denial Of Service (DoS)

DoS attack [5] main intention is to make the machine or network not available to the users and it comes under the category of active attack. DoS attack interrupts the work of the intended users which in turn wastes their time and resources and cause loss. There are mainly two types of DoS attacks. Buffer overflow attack and Flood attack are the two types of attacks. In Buffer overflow attack the malicious actor causes the machine's all the available RAM and hard disk space to be consumed which can lead to denial of service to the legitimate users. In case of flood attack , the attacker should have higher bandwidth than the target machine to saturate it with overwhelming number of packets which results in denial of service. General indicators of this attack can be slow functioning of the network and slow loading of any website you requested etc. Some famous DoS attacks are Pod, Smurf attack, teardrop.

b. Probe Attack

In the probe attack, the attacker have an access to the information about the whole network before launching an attack. Examples: ipsweep, nmap, portsweep, satan attacks.Ipsweep which is also known as ICMP sweep attacks. Attacker sends ICMP echo requests to multiple destination addresses. If the user who receives them, replies to these requests then it reveals the targets IP address to the attacker. The primary purpose of the IP sweep is to determine which ranges of Internet Protocol (IP) addresses correspond to active hosts. In port sweep, the attacker scans the same port on several computers to see if the user is active on specific port. If the user is active the attacker sends the ICMP echo requests.

DOS	Probe	U2R	R2L
apache2	saint	imap	perl
land	mscan	xsnop	sqlattack
back	ipsweep	multihop	worm
Neptune	nmap	warezclient	loadmodule
mailbomb	satan	spy	ps
teardrop	portsweep	named	rootkit
smurf		xlock	bufferoverflow
udpstorm		phf	xterm
pod		snmpgetattack	snmpguess
		warezmaster	
		ftp_write	
		httpunnel	
		guesspasswd	
		sendmail	

Figure 2. Representation of Attack types

c. Root to User(R2L) Attack

By sending packets to a remote machine, the attacker takes advantage of various network vulnerabilities and acquire local access. Examples: Imap, guess password and ftpwrite, spy, multihop, warez client, warez master attacks. MAP is used to transmit username and password of the messages. It also sends the contents of the message. In imap the user tries to access the credentials of the user and tries to access information. Warez client and the Warez master exploits the vulnerabilities present in anonymous FTP. These attacks may occur both in the Linux and windows systems.

C. Dataset description

For this deep learning model, we have chosen the NSL-KDD dataset [6], is a benchmark dataset for training and validating the model. The most important features from the dataset are services, protocols and flags. Protocols used in the dataset are TCP, UDP and ICMP. Some of the services are telnet, http, smtp to a total of 70 services. Because the NSL-KDD dataset doesn't contain duplicate items in the training dataset, and the classifiers are not slanted towards more often records. It is recommended to include the duplicates in the test sets if there are any; techniques with higher detection rates on frequent records have no influence on learners' outcomes.

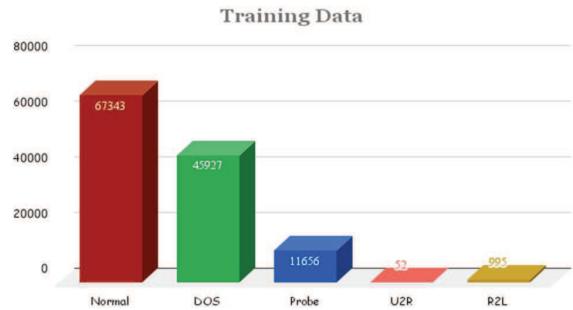


Figure 3. Representation of NSL - KDD Training Data set

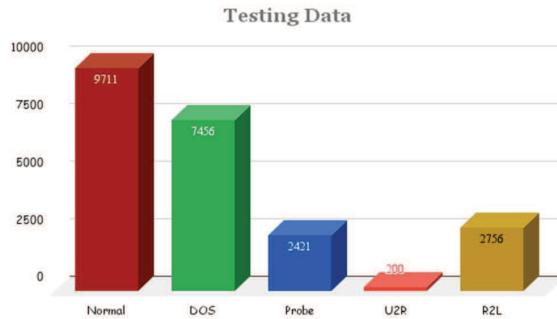


Figure 4. Representation of NSL - KDD Testing Data set

II. LITERATURE SURVEY

This section describes the literature review that were utilized as references. Eric Marcell Jones Jr., Sara A Althubiti, and Kaushik Roy [7] discussed that flow based CIDDS-001 data was already analyzed using variety of machine learning models and they are currently implementing LSTM model for deep by using Keras which is a high- level Application Programming Interface using Python. The entire dataset is splitted as 67% for training purpose and 33% for validation purpose. LSTM model is trained with 200 epochs and after several observations the learning rate was adjusted as 0.01. “Rmsprop” optimizer was applied as it is more apt for larger datasets like CIDDS. In the end they concluded that rmsprop optimizer is more suitable for LSTM model and their model has reached an accuracy of 0.87.

Chongzen Zhang, Fangming Ruan et al. [8] discussed the methodology of taking encoder of autoencoder benefit to construct a network , simple network structure can make prediction quickly and is an exploration of a large flow modern network . The experiment is evaluating the classification model performance within 5 classifications(DoS, Normal, Probe, R2L, U2R). Possibility of false alarm and false negative can be reduced with the proposed method.This model is more effective when compared to other similar deep learning models.Due to limited samples available for training, detection rate of R2L and U2R attacks within five classes is low , this reduces the overall accuracy of the NIDS.

Pengcheng Wang, Jun Lin, Lan Liu, Langzhou Liu [9] published this paper. Used the CSE-CIC-IDS2018 and NSL-KDD datasets for intrusion detection. As the network traffic is highly imbalanced, they have proposed DSSTE algorithm for compressing the majority samples and thereby reducing the variation in the dataset. They have used several classifiers like Random forest, LSTM, Mini- VGGNet, SVM, Random Forest and Alex Net for classification models. They have concluded that the high imbalances in the datasets lead to improper prediction of malicious attacks.

Nongmeikapam Brajabidhu Singh, Arindam Sarkar, Moirangthem Marjit Singh, Jyotsna Kumar Mandal [10] has Used the CSE-CIC-IDS2018 and NSL- KDD datasets for intrusion detection. As the network traffics highly imbalanced, they have proposed DSSTE algorithm for compressing the majority samples and thereby reducing the variation in the dataset. They have used several classifierslike Random forest, LSTM ,Mini-VGGNet , SVM, Random Forest and AlexNet for classification models. They have concluded that the high imbalances in the datasets lead to improper prediction of malicious attacks.

M. Sood and N. Bindra [11] had used Anomaly detection is performed using Machine Learning techniques, which were trained to recognize abnormal activity. They have mainly focused on detecting DDoS attacks since they are most frequent attacks and they need to be detected immediately in order to decrease the loss of the organization and testing for DDoS is challenging. They inspected the performances of five supervised machine learning classifiers which KNN, Random Forest, GNB, SVM and Logistic Regression. Their conclusions were that Random forest can be chosen for DDoS detection as it is a better choice and it had achieved an accuracy of 96%.

Algorithm

Input : Training dataset of NSL-KDD dataset

Output : Classifying attacks on Testing set of NSL-KDD dataset

1: Step – 1 : Data Exploration, analyze andunderstand the dataset

2: Add the column names to both the training andtesting dataset

3: Step – 2 : Label Aggregation

4: Identify the unique types of attacks for both training and testing dataset

5: Add a column with the name class with the different class of attacks like DoS, Probe, R2L

6: Step – 3 : Data Preprocessing, transform data into some useful information

7: Classify the features into object types and non-objectforms

8: Using OneHotEncoding the object type data is converted to numerical data with the help of dummy variables.

9: Minmax Scalar transforms the features to a specific range.

10: Feature Extraction is done using Autoencoder(encoder, decoder)

11: Step – 4 : Training the Data, train using the Recurrent Layers

12: The model is Sequential with LSTM, GRU, and Dense layers as the output.

13: Fit the model with the batch size, epoch values, andcallback as Early Stopping and mode is minimum

14: Step – 5: Testing or Evaluating the Performance

15: Predict the outcome with the classification report, confusion matrix and multi-label matrix

III . PROPOSED METHODOLOGY

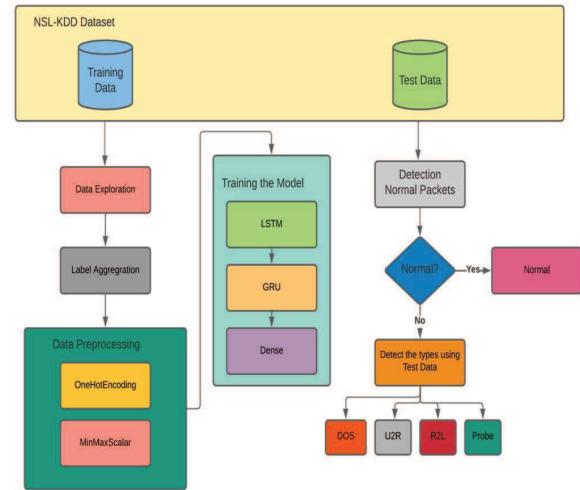
The flow chart for the proposed methodology describes the processes of data acquisition, data preprocessing, feature extraction using Autoencoder [12] , Training the model with different algorithms and testing the model. The process flow diagram gives a sequence of actions taking place in the proposed methodology. The process flow diagram is as shown in Fig 3 describes various modules like data pre-processing, training the model and testing the model. If the accuracy is unsatisfactory, we can again train the model with different values of epochs and changing of functions.

In order to iteratively process the data, the deep learning is very helpful. We randomly initialize the values of the weights and biases and work towards the right values by trial and error. The model makes use of the both parameters and hyper parameters. The parameters include weights and biases, whereas the hyper parameters include the total number of layers in the model, nodes in each layer, optimizers, activation function, cost function, batch size and epochs. The validation of the data set plays a very important role, The model is predicting on the validation data set. The values of accuracy and loss are calculated and the model is fine-tuned and the learning process is repeated on the results. When we fit the model we use arguments like batch size and epoch. The batch represents the set of the samples sent through the model in a single pass. The training data is divided into one or more batches. The epoch represents the number of times the entire training sent through the model. Generally, it is believed that the increased number of epochs better the accuracy, but it is not always true. Sometimes the model can over fit, then we need to reduce overfitting of the data.

The proposed methodology comprises of three modules. In the first module we perform data exploration ,label aggregation. When we import the dataset, we get the files in text format and we need to convert them in to .csv format in order to perform the next steps. Initially the csv file doesn't contain column names. So, we need to add column names for both training and testing datasets. We explore the data on protocol types, services it consists, the outcome of the training and testing datasets to classify them into different types of attacks. In label aggregation we check the unique outcome of the training and testing datasets and we classify these outcomes according to the type of attacks like Dos , Probe, R2L. We add a column name class and give the categories for different types of attacks existed in the dataset like buffer overflow , rootkit, worm, SqlAttack etc. We give the labels to the attacks. In the second module, we implement data preprocessing and the feature extraction with Auto Encoder. We use MinMaxScaler and OneHotEncoder [13] in data preprocessing . The One Hot Encoding is used to convert the nominal data into numeric. It is represented in the form of binary variables which are also known as dummy variables. The MinMax scaler is used to transform the features by scaling each feature to a specific or to a given values of particular range. Each feature is generally scaled between 0 to 1. The feature extraction is done with AutoEncoder. The AutoEncoder consist of three main important components

encoder, code and decoder.

The AutoEncoder is used for dimensionality reduction. The encoder compresses the input and produce code and the decoder reassemble the input using the code. It ignores the noise in the data. In this methodology Early stopping is used to avoid Overfitting of data. The third module comprises of training and testing the module for performance. We use Gated Recurrent Unit to train the model. GRU is considered to give more efficient and it is getting popular for the sequential data models. The GRU solves the vanishing gradient problem and it remembers the information for long period of time. This model uses soft-max as an activation function, sparse categorical loss entropy and Adam Optimizer [14] and the model fit is done after reshaping the train data with 50 epochs. Testing the model for performance is done.



The model's performance consists of precision, recall, accuracy, and f1 score. We can determine them using true positives, false negative, false positives, true negatives, and values. A true positive value is when our model accurately predicts positive class and when a model correctly predicts the negative class, it is said to be True Negative. When the model predicts the negative class incorrectly, it generates a false negative. When the model predicts the positive class of variables incorrectly, the result is False Positive. We employ a multi label confusion matrix to gather two-dimensional data on true and false positives and negatives. The multi-dimensional data is subsequently converted into a contiguous flattened array using the ravel technique.

To simplify the equations let us consider true positive as P, false positive as Q, true negative as R, false negative as S.

$$Accuracy = \frac{P + R}{P + Q + R + S} \quad (1)$$

$$Precision = \frac{P}{P + Q} \quad (2)$$

$$Recall = \frac{P}{P + S} \quad (3)$$

$$f1_score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

The values of precision, f1_score, accuracy, and recall values are rounding up to 4 places and the results of the performance are shown in Results and Analysis.

IV. RESULT AND ANALYSIS

In the increasing order of the accuracies of the different attacks. The Probe exhibits high accuracy among all other types of attacks, about 95%, and the next highest accuracy is exhibited for the Probe attack of 92%, followed by R2L, which is 87% and normal attacks with 77%. Precision is used to calculate the number of labels predictions that truly belong to labels of the class. The has the highest precision, it predicts the testing samples with precision of 0.96.

	accuracy	precision	recall	f1_score
DoS	0.9253	0.9571	0.8113	0.8782
Probe	0.9651	0.8777	0.7856	0.8291
R2L	0.8894	0.9444	0.1473	0.2549
normal	0.8053	0.6971	0.9717	0.8118

Figure 6. Output results for the implementation

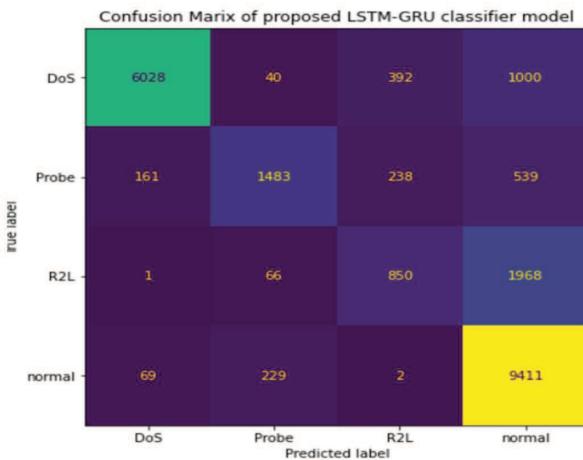


Figure 7. Confusion Matrix

V. CONCLUSION AND FUTURE WORK

The suggested approach utilizes both LSTM and GRU. It is useful for sequence modelling techniques .A Gated Recurrent Unit is a variant of the RNN design and employs a gated process to control and manage the flow of information between cells in the neural networks.

We implemented NIDS with GRU and also used Dense layer. For now, the we didn't get much higher accuracy . So, we will try to increase the accuracy of the algorithm with NSL- KDD dataset .And then implement GRU for different large datasets like CICIDS 2017 , CSE-CIC-IDS 2018 [15] dataset and UNSW_NB 15. Implement the Intrusion Detection system with more advanced versions of Gated Recurrent Unit and other deep learning techniques available. Among the research community, there is a growing desire for a trustworthy and real-world attacks dataset. It is feasible to work in multiple ways in the future, one of which is to improve on the existing work by implementing alternative balancing algorithms and investigating their effects on model learning.

REFERENCES

- [1] Y. B. Y. G. LeCun, "Deep learning," *Nature*, 2015.
- [2] W. a. S. I. a. V. O. Zaremba, "Recurrent Neural Network Regularization," *arXiv preprint*, 2014.
- [3] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.
- [4] R. a. S. F. M. Dey, "Gate-variants of Gated Recurrent Unit (GRU) neural networks," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2017, pp. 1597-1600.
- [5] T. a. X. Y. a. S. G. a. J. W. Mahjabin, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques," *International Journal of Distributed Sensor Networks*, vol. 13, p. 1550147717741463, 2017.
- [6] "UNB," Canadian Institute of CyberSecurity, [Online]. Available: <https://www.unb.ca/cic/datasets/nsi.html>. [Accessed 15 12 2021].
- [7] S. A. Althubiti, E. M. Jones and K. Roy, "LSTM for Anomaly-Based Network Intrusion Detection," in *2018 28th International Telecommunication Networks and Applications Conference (ITNAC)*, Sydney, NSW, Australia, 2018.
- [8] C. Zhang, F. Ruan, L. Yin, X. Chen, L. Zhai and F. Liu, "A Deep Learning Approach for Network Intrusion Detection Based on NSL-KDD Dataset," in *2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, Xiamen, China, 2019.
- [9] L. Liu, P. Wang, J. Lin and L. Liu, "Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning," *IEEE Access*, vol. vol. 9, no. 2021, pp. 7550-7563, 2021.
- [10] N. B. M. KumarMandalc, "A novel wide & deep transfer learning stacked GRU framework for network intrusion detection," *Journal of Information Security and Applications*, vol. 61, 2021.
- [11] M. Sood and N. Bindra, "Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset," *Automatic Control and Computer Sciences*, no. September 2019, p. 419–428, 2019.
- [12] Y. N. a. N. S. a. S. D. a. Z. A. a. o. Kunang, "Automatic features extraction using autoencoder in intrusion detection system," in *2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2018, pp. 219-224.
- [13] M. A. B. J. G. S. E. Pau Rodríguez, "Beyond one-hot encoding: Lower dimensional target embedding," *Image and Vision Computing*, vol. 75, pp. 21–31, 2018.

[14] Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, 2018, pp. 1-2.

[15] J. L. , T. M. Leevy, "A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 Big Data," *Journal of Big Data*, vol. 7.

Diagnosing Musculoskeletal Disorders from Shoulder Radiographs using Deep Learning Models

A. Kavitha

Vel Tech Rangarajan Dr Sagunthala
R&D Institute of Science and
Technology
Chennai, Tamilnadu, India
akavitha@veltech.edu.in

T. Kujani

Vel Tech Rangarajan Dr Sagunthala
R&D Institute of Science and
Technology
Chennai, Tamilnadu, India
kujani@veltech.edu.in,

S. Gokul Kannan

Vel Tech Rangarajan Dr Sagunthala
R&D Institute of Science and
Technology
Chennai, Tamilnadu, India
drsgokulkannan@veltech.edu.in

V. Akila

Gokaraju Rangaraju Institute of
Engineering and Technology,
Hyderabad, Telangana, India.
akila1584@grietcollege.com

Abstract— In recent years, exploring how Artificial Intelligence, especially deep learning algorithms, may be used to analyse healthcare data has become crucial. Deep learning can be employed for assisting the doctor and supplementing the process of diagnosing diseases. Early diagnosing of the problems of the shoulder, a vital body part will be useful to manage or solve the shoulder pain with efficacy. Our study has utilized deep learning to diagnose abnormalities from radiographs of the upper extremity, specifically the shoulder. The abnormality in the shoulder has been detected from the radiographs of the shoulder using four pre-trained Convolutional neural networks (CNN) models such as VGG-16, DenseNet-201, ResNet-50, and EfficientNet-B1. We have also calculated the accuracy, precision, recall, specificity, sensitivity, F1score, and area under the receiver operating characteristic (AUROC). Shoulder radiographs from the MURA dataset were used to test the four CNN models. From the experiments conducted, both ResNet-50 and VGG-16 had the maximum accuracy of 81%; ResNet50 had the best AUC score of 0.74 and it outperformed the other three models.

Keywords—*Musculoskeletal, Deep learning, detection, Convolutional Neural Network, CNN, EfficientNet, VGG, DenseNet, ResNet.*

I. INTRODUCTION

Biology, physics, medicine, and engineering all play a role in biomedical image analysis. It is concerned with the use of image processing techniques and machine learning approaches to solve biological or medical issues. Radiographical photographs to be examined include a wealth of information about the anatomical structure under inquiry, allowing clinicians to make accurate diagnoses and hence choose appropriate treatment. These medical photos are manually investigated by doctors using visual interpretation.

According to a recent examination of Global Burden of Disease (GBD) data, musculoskeletal problems affect nearly 1.71 billion people globally, and they are the leading cause of severe, long-term pain and impairment. Musculoskeletal problems affect movement and dexterity severely, resulting in early retirement from work, poor well-being, and a reduced ability to participate in society. As a result of population expansion and ageing, the number of people disabled by musculoskeletal illnesses has been increasing in recent decades, and this trend is projected to continue.

According to a study [4], radiologist become fatigued at the end of the day due to eye/visual strain; the accuracy of

diagnosing the abnormality in the radiographs is being reduced. Hence, automizing the detection of abnormality will increase early treatment for the abnormal patients. It is tedious to identify the abnormality in a radiographical images of human body. Various techniques in machine learning techniques have been utilized to detect the abnormality in the radiographical images.

Using the publicly accessible MURA [1] musculoskeletal dataset, Banga and Waiganjo [5] have created ensemble200 model to detect abnormalities. The models were created using a combination of theoretical research, iterative prototyping, and empirical testing. The model outperforms the DenseNet model in terms of F1 score.

Amila et al. [6] have presented an Artificial Neural Network (ANN) model for diagnosing and predicting osteoporosis by combining risk factors, metabolic testing, and x-ray scans. Age, sex, fracture details, glucocorticoids use, T-score, smoking, vitamin D, and calcium were all employed in this study to predict osteoporosis bone disease.

Saif et al. [7] have developed a capsule network design for musculoskeletal radiograph abnormality detection, and this CapsNet architecture has proved to have extremely promising qualities that can help overcome Convolutional Neural Network (CNN)'s limitations. Furthermore, this capsule network has a kappa score that is 10% greater than a Densenet169.

The pre-trained CNN models such as VGG-19, ResNet-50, EfficientNet-B1 and DenseNet-201 have employed for the diagnosing the shoulder problems with MURA dataset. The predicting accuracy of the approaches based on the various performance metrics such as accuracy, precision, recall, F1-score, and AUC score have been analysed. It is anticipated that this research help to diagnose at the expert level, hence enhancing healthcare access in areas where professional radiologists are scarce.

II. MATERIALS AND METHODS

A. Study Design

The main goal of this study is to use shoulder radiographs to assess the likelihood of musculoskeletal abnormalities. To diagnose abnormalities from shoulder radiographs, the various pre-trained CNN models have been used. The accuracy of the diagnosis of abnormalities and several CNN models were studied.

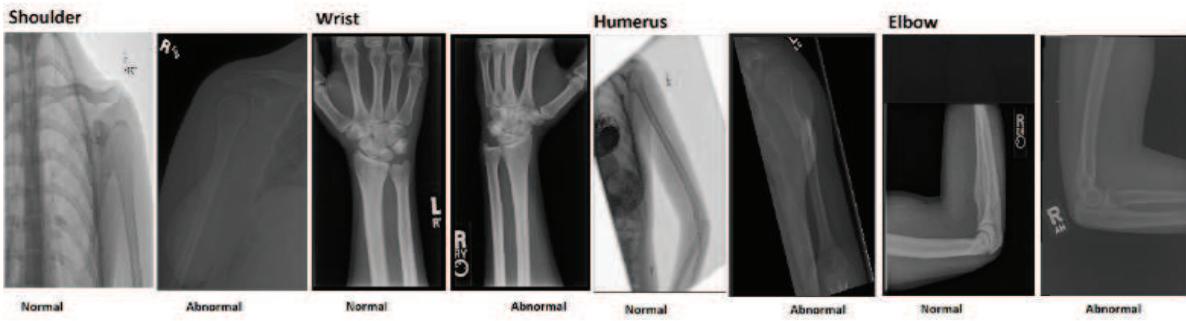


Fig. 1. Sample radiograph images of MURA dataset

B. Data Acquisition

The data set on musculoskeletal radiographs from Stanford ML Group [1,2] has been used in this paper. MURA (musculoskeletal radiographs) is a vast database of bone X-rays. The responsibility of assessing whether an X-ray scan is normal or abnormal is delegated to algorithms. MURA is the most comprehensive public radiography image databases available. MURA is a musculoskeletal radiograph dataset containing 40,561 multi-view radiological images of upper extremities in total. It contains 14,863 investigations from 12,173 individuals. The seven standard upper extremities radiography study types are the shoulder, elbow, finger, forearm, hand, humerus, and wrist. The shoulder dataset which consists of 8743 images have been employed for the proposed work. Sample radiograph images of the MURA dataset have been illustrated in Fig.1 which shows both normal and abnormal images of shoulder, wrist, humerus and elbow.

C. Method - CNN Model Architecture

Machine learning is a competent technique in decision making in healthcare industry. The computational approach CNN has been used for diagnosing the problems in the shoulder. It was chosen since they are the most effective segmentation and object detection models. LeCun et al. (2015), Krizhevsky et al. (2012).

CNN is a class of Deep neural network, breakthrough image identification. The CNN is a Feed Forward Neural Network that is used to evaluate images by processing data using a grid-like topology. ConvNet is another name for it. The CNN's main advantage is that it demands less effort. Other algorithms use more pre-processing than this. The CNN's potential is that it can capture the reliance of picture spatial and temporal information for better understanding comprehending the image. The ConvNet is used to reduce the size of the image so that it can be processed quickly. The basic operation used in CNN is convolution operation. It takes input images and passes through various layers like input, hidden and output layer and produces the output as name of image. Input layer takes pixels in the input image and arrange in the form of arrays using pixel values.

Hidden layers extract features of images called as feature extraction by manipulation and calculation procedure. There is various level in the hidden layer are convolution layer, ReLU layer and pooling layer. Convolution layer uses the matrix filter to detect the pattern in the image. It means twist the data or alter it. Convolution operation in

the convolution layer is performed by using various filters. Sliding the filter through the input matrix and using a dot product to generate a convolved feature that is the same size as the matrix filter.

The next is Rectified Linear Unit (ReLU) layer which is the activation function used in the neural network. It works each and every element. This layer makes all negative pixels to zero. The output is rectified feature by using multiple filters.

The pooling layer extracts various portions of the images, such as corners and beaks. The rectified feature map is then put through a pooling layer, yielding a pooled feature map with the most features. The process of transforming a two-dimensional pooling feature map into a continuous single-dimensional vector is known as flattening.

The output layer is a fully linked layer that recognizes the object from the flattened matrix as an image. The pixel from the flattened matrix is provided as input to the fully linked network's various layers. The output layer is where the image classification takes place. The architecture of a Convolutional Neural Network is shown in Figure 2.

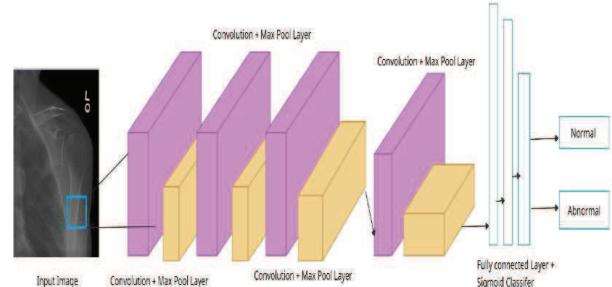


Fig. 2 Architecture of Convolutional Neural Network

VGG16: VGG16 (Visual Geometry Group) was presented by Simonyan et al. (2014); it is a simple and widely used Convolutional Neural Network (CNN) Architecture. In this architecture, large size kernel filters such as 11 and 5 are replaced with many 3*3 kernel filters. Its very helpful to learn complex features at low cost. This architecture utilizes 16 layers. The architecture of VGG-16 is depicted in Fig.3.

DenseNet-201: Huang et al. (2017) have come up with Dense CNN (DenseNet). This creates a feed-forward connection between each layer and the others. It has direct connections of $L(L+1)/2$. Each layer uses the feature-maps of all previous levels as inputs, and its own feature-maps

are utilised as inputs into all following layers. It can be scaled to hundreds of layers. Denseblock with 5 layers is shown in Fig. 4 (a); deep DenseNet with three dense blocks is shown Fig. 4 (b).

ResNet-50: One important problem to be addressed in deep layers of CNN is sometimes there may be more error as the number of layers are increased and also to handle the problem of vanishing gradients. To overcome this issue, He et al. (2016) have introduced the deep residual learning framework which uses skip connection technique for skipping the training from few layers and directly connecting to output. The potential part in Residual Network is that re-formulating the layers as residual functions with reference to the input layer. In Resnet there are two kinds of blocks, one is identity block and other is convolution block. When the input to the network and output of the network is same, then identity block is used. Convolution block is used when the inputs size and output size is not same. Fig. 5. shows the architecture of ResNet-50.

EfficientNet-B1: Tan et al. (2019) have presented EfficientNet utilising effective compound scaling method. It scales up the baseline ConvNet to any target requirements by scaling the depth by 20%, width by 10%, and resolution by 15%. Fig. 6. shows compound scaling method of EfficientNet-B1 architecture.

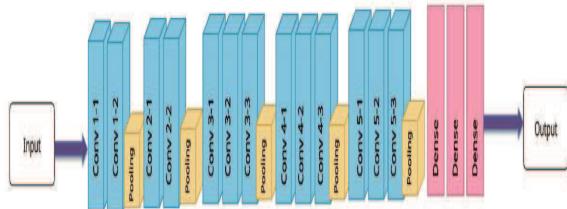


Fig. 3. VGG – 16

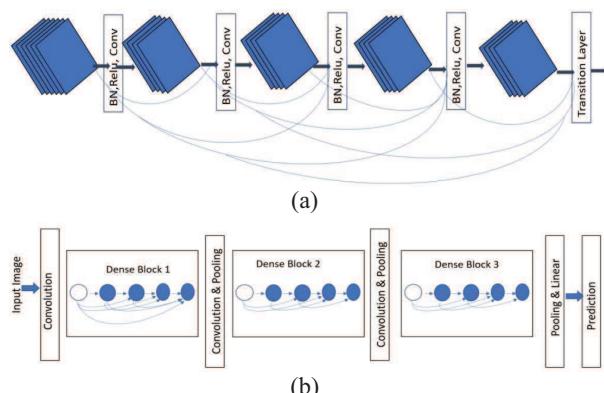


Fig. 4. (a) Denseblock with 5 layers (b) A deep DenseNet with three dense blocks

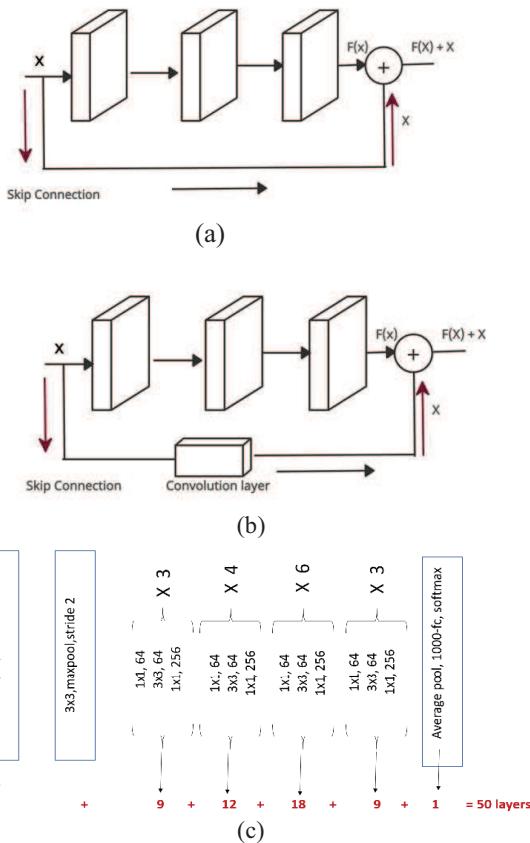


Fig. 5. (a) Identity Block (b) Convolution Block (c) ResNet-50 Architecture

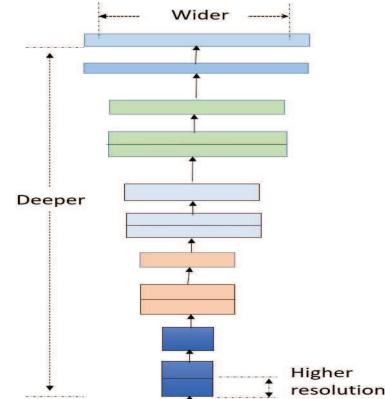


Fig. 6. Compound Scaling Method (EfficientNet-B1)

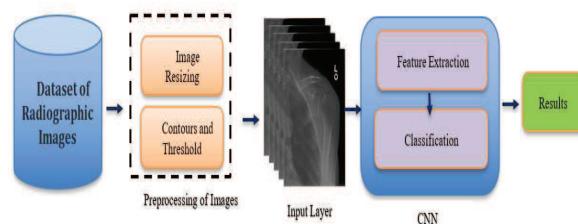


Fig. 7. Overall block diagram of the proposed approach

III. IMPLEMENTATION DETAILS

A pre-trained model decreases on the amount of time and data needed to build a reliable classifier. Hence, the pretrained models such as VGG-16, DenseNet-201, ResNet-50, and EfficientNet-B1 were tested for evaluating their performance on MURA's shoulder dataset.

Data augmentation is a process by which the quantity of images is increased tremendously to avoid overfitting in the deep learning models. During the preprocessing of images, images were augmented by flipping the images horizontally. The images are rescaled to values between 0 and 1. The images are rotated to a maximum of 30 degrees with width/height shift range of 0.3. The horizontal flip of the images is also permitted. For scaling of the image, the

nearest neighbor algorithm is used. Data augmentation is carried out only for the training set. For the testing data, the images are only rescaled to values between 0 and 1.

The target size of the feature maps at the end of CNN model is 64 x 64. For the training set, while performing the back-propagation algorithm the batch size is fixed at 64 and shuffling of the images is performed in each iteration, whereas, for the testing dataset the batch size is 1 and shuffling is not permitted.

The parameters for the pre-trained models, data augmentation and data generator are chosen as per the experiments conducted and are shown in Table 1 and Table 2. Fig. 7. shows overall block diagram of the proposed approach.

Table 1. Parameters of Pre-trained Models

Model Parameters /Model name	VGG16	DenseNet201	ResNet50	EfficientNetB1
include_top	FALSE	FALSE	FALSE	FALSE
Input	64,64,3	64,64,3	64,64,3	64,64,3
weights	imagenet	imagenet	imagenet	imagenet
pooling	avg	avg	avg	avg
classes	2	2	2	2
Output shape	512	(None, 1920)	(None, 2048)	(None, 1280)
Number of parameters	14714688	18321984	23587712	6575239
activation	softmax	softmax	softmax	softmax
optimizer	RMSprop(lr=0.0001)	RMSprop(lr=0.0001)	RMSprop(lr=0.0001)	RMSprop(lr=0.0001)
loss	binary_crossentropy	binary_crossentropy	binary_crossentropy	binary_crossentropy

Table 2. Parameters for data augmentation and data generator

Data Augmentation		Data Generator	
Parameters	Value	Parameters	Value
Training data		Training data	
rescale	1./255	target_size	64,64
rotation_range	30	batch_size	64
width shift range	0.3	class_mode	Categorical
height shift range	0.3	shuffle	TRUE
horizontal_flip	TRUE	target_size	64,64
fill_mode	nearest	Test data	
rescale	1./255	target_size	64,64
Test data		class_mode	Categorical
rescale	1./255	shuffle	FALSE

III. RESULTS AND DISCUSSION

The experiments were conducted with the dataset of shoulder. Training dataset consist of 4113 negative and 4087 positive radiographic images. The proposed method has been comprehensively evaluated on the validation dataset which consists of 262 positive and 281 negative

radiographic images. Totally 8743 images have been used for the experiment. For experimenting, we have explored TensorFlow framework to implement the network model and performed training on shoulder dataset. Programming Language python has been used. Performance metrics have been discussed in the following subsection.

A. Performance Metrics

The following performance indicators were used to evaluate the CNN models employed in the abnormality detection (1) Accuracy (2) Precision (positive and negative samples), (3) recall (positive and negative samples), (4) F1 score, and (5) AUC score are all factors to consider. In the following equations, TP measures the number of true positive results and TN measures the number of true negative results. FP indicates the number of false positives and the FN indicates the number of false negative results.

- Accuracy - simply the count of correct decisions divided by the total number of decisions

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision is the proportion of predictions that are actually true. It is calculated separately for the positive and the negative samples. For positive samples, the precision is given with the letter “p” in brackets. Similarly, for the negative samples the letter “n” is used.

$$\text{Precision } (p) = \frac{TP}{TP + FP}$$

$$\text{Precision } (n) = \frac{TN}{TN + FN}$$

- Recall is the proportion of correct predictions for the given data set. It is also calculated separately for the positive and the negative samples.

$$\text{Recall } (p) = \frac{TP}{TP + FN}$$

$$\text{Recall } (n) = \frac{TN}{TN + FP}$$

- The harmonic mean of precision and recall at a particular position is the F-measure.

$$F - \text{measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

- AUC-ROC is the area under the ROC curve and ranges between 0 and 1. ROC curve is plotted for different classification thresholds using recall and false positive rate (i.e., 1 - specificity). AUC score has also been calculated.

From the experiments conducted, it observed that maximum accuracy of 0.74 has been achieved with ResNet-50 and VGG-16 and is depicted in Fig. 8. The best precision for positive classes and negative classes of 0.95 and 0.85 and are achieved by VGG-16 and ResNet-50, respectively. The best F1-score for positive classes and negative classes of 0.8 and 0.87 are achieved by VGG-16 and ResNet-50 respectively. Fig. 9. shows the results of the performance evaluation of the four CNN models using radiographs of the shoulder. The AUC score also was estimated; AUC is also plotted and is depicted in Fig. 10. ResNet-50 had the highest AUC score of 0.74.

Accuracy of Transfer Learning Models

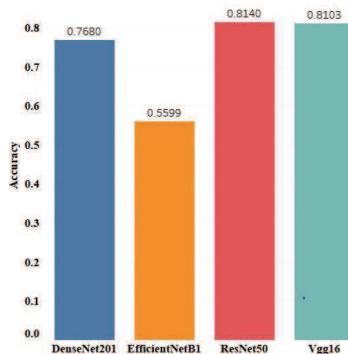
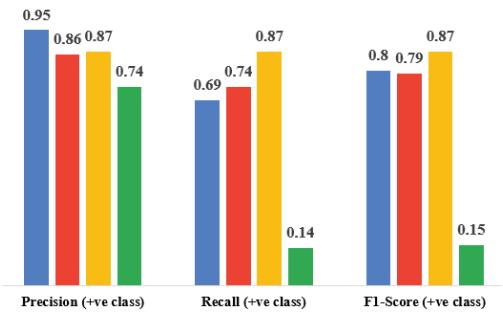
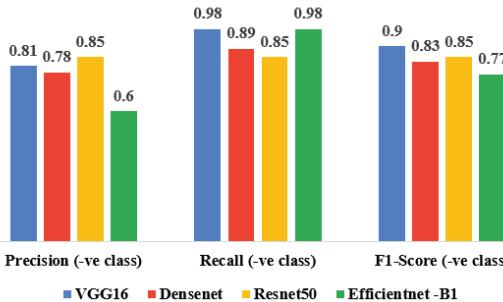


Fig.8. Accuracy of CNN based models (Densenet-201, EfficientNet-B1, ResNet-50, VGG16)



(a)Positive class



(a) Negative class

Fig. 9. Performance Comparison of DenseNet-201, EfficientNet-B1, ResNet-50, VGG-16 models

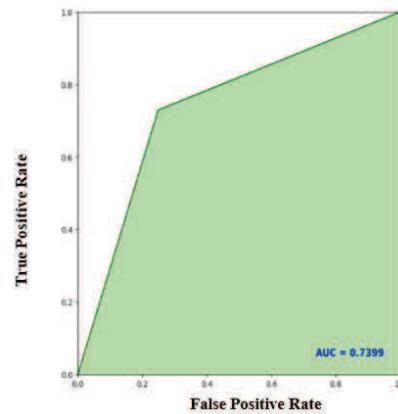


Fig. 10. AUC of ResNet-50

IV. Conclusion

The deep CNN model has been utilised for accurately diagnosing the abnormality from the radiographs of shoulder. The proposed approach used CNN which takes the input from data set of radiographical images of shoulder. Several experiments have been conducted on the MURA data set with various predefined CNN models and analysed the performance of CNN models. Among the CNN models which were experimented, ResNet-50 had the highest AUC score of 0.74; VGG-16 and ResNet-50 had the same accuracy of 81%. These findings may have a significant impact on diagnosing the musculoskeletal problems in the medical industry. In future, research work may be extended with custom Deep CNN model to improve the accuracy of abnormality detection.

ACKNOWLEDGEMENT

We thank the Stanford Program for Artificial Intelligence in Medicine and Imaging (AIMI.stanford.edu) for clinical dataset infrastructure support.

REFERENCES

1. Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R.L. and Langlotz, C., 2017. Mura: Large dataset for abnormality detection in musculoskeletal radiographs. *arXiv preprint arXiv:1712.06957*.
2. <https://stanfordnlggroup.github.io/competitions/mura/>
3. <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions>
4. Krupinski, E.A., Berbaum, K.S., Caldwell, R.T., Schatz, K.M. and Kim, J., 2010. Long radiology workdays reduce detection and accommodation accuracy. *Journal of the American College of Radiology*, 7(9), pp.698-704.
5. Banga, D. and Waiganjo, P., 2019. Abnormality detection in musculoskeletal radiographs with convolutional neural networks (ensembles) and performance optimization. *arXiv preprint arXiv:1908.02170*.
6. Amila P., Nejra P., Nejra P., Adis P., Matej P., Amina P. (2021) Using Artificial Intelligence in Prediction of Osteoporosis. In: Badnjevic A., Gurbeta Pokvić L. (eds) CMBEBIH 2021. CMBEBIH 2021. IFMBE Proceedings, vol 84. Springer, Cham. https://doi.org/10.1007/978-3-030-73909-6_33
7. Saif, A.F.M., Shahnaz, C., Zhu, W.P. and Ahmad, M.O., 2019. Abnormality detection in musculoskeletal radiographs using capsule network. *IEEE Access*, 7, pp.81494-81503.
8. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, pp.1097-1105.
9. Rizzoli, R., Body, J.J., Brandi, M.L., Cannata-Andia, J., Chappard, D., El Maghraoui, A., Glüer, C.C., Kendler, D., Napoli, N., Papaioannou, A. and Pierroz, D.D., 2013. Cancer-associated bone disease. *Osteoporosis International*, 24(12), pp.2929-2953.
10. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
11. Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708. 2017.
12. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
13. Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In International Conference on Machine Learning, pp. 6105-6114. PMLR, 2019.
14. <https://neurohive.io/en/popular-networks/vgg16/>
15. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436-444.

Piezoelectric Transducers Power Generation for Low Power Applications

B.Rubini

*Department of Electrical and Electronics Engineering
Vels institute of science technology and advanced studies
Chennai
India.
Rubini.se@velsuniv.ac.in*

S.Manoj

*Department of Electrical and Electronics Engineering
Vels institute of science technology and advanced
studies Chennai India.
Manoj.se@velsuniv.ac.in*

Abstract— The demand for energy is increasing on a daily basis. The utilization of waste energy as reusable energy of human foot power is very relevant and important. On the other hand, many energy resources are being depleted and squandered. It will generate electrical force when people walk around it. The power generation of piezoelectric transducers is organized in a planned circuit alongside various parts that are connected in series under a mechanical design. The observing circuit is constrained by a customized Arduino Uno. These electrical charges will be stored in a super capacitor and can be used in low-power applications. The value of the DC voltmeter will be displayed on the LCD. This experiment yielded the most cost-effective solution.

Keywords—Piezo transducer, Super capacitor, Pressing mechanism, low Power Absorption.

I. INTRODUCTION

Since the invention of the human being, the use of electrical energy has grown at a rapid pace. As human populations have declined over time, electricity consumers have been content to rely on traditional energy sources. Later, as the population grew, so did the demand for energy, so we developed non-conventional energy sources, but these methods are no longer completely meeting the demand. If such strategies are used, there are a variety of techniques for delivering power. Powerage piezoelectric transducers can be a compelling strategy for producing power or capacity for advances in low force electronic applications. Strolling is a common activity in human life. When a person walks, he loses energy on the street surface in the form of vibration, sound, and so on, as a result of the exchange of his weight onto the street surface through footfalls on the ground during each progression. This energy can be converted into a usable structure, such as an electrical structure. Power has become a lifeline for the human population. Power is becoming increasingly important. To perform various tasks, some innovations necessitate a large amount of electrical ability. As we all know, power is generated by various sources such as water, wind, and so on. We can measure the voltage generated

by connecting a voltage sensor to an Arduino Uno, which converts the analogue signal to a digital signal and interfaces with a 16x2 LCD display. A power measuring metre, also known as a multi meter, is used to measure millivolts and micro-amps. Its measuring value is accuracy. Nowadays, the most popular method of power generation is thermal power generation (76.299MW) (52 percent). This process burns coal to produce steam, which pollutes the environment, causes land loss, and is very expensive.

II. OBJECTIVE

The main goal of these activities is to foster a successful method of force age technique. The goal of this piezoelectric transducers power age project is to change over stride, attempting to convert strolling and running energy into electrical energy. It is used to generate electricity even while striding. The supply of electrical energy is rapidly expanding. In any case, power age ordinary assets are insufficient for an absolute interest in electrical energy. As a result, many analysts are dealing with non-regular methods of electrical force age. Piezoelectric transducers power age framework is also a non-traditional energy creation framework. It converts the mechanical energy of strides into electrical energy by using transducers. This force age framework has the potential to become extremely well known among nations such as Pakistan, China, and India.

- Power generation simply walking on step.
- To real time dc voltage monitoring system.
- It can be implemented on road, bus station, many public places, dance Floors and flooring tiles.

Piezoelectric transducers arrangement board: It has a total of 30 piezoelectric transducers; 25 in parallel and 5 in series, with a bridge rectifier and a capacitor on each transducer. The bridge rectifier converts alternating current to direct current, and the capacitor receives pure direct current output. Red (+) and black (18awg wire) were used (-). Voltage controllers 7805 are to be used to

achieve these results. Charge controller: 1000uf capacitor can store voltage from direct piezoelectric board. These controllers can provide nearby on-card guidance, which eliminates the appropriation issues associated with single point guidance and 330uf capacitor flow the electric charges to Super capacitor. 1.5F 5.5V super capacitor In comparison to a rechargeable battery, a super capacitor charges and discharges quickly.

A super capacitor is also a rechargeable battery. Measuring Meter (voltage & current): This metre is used to measure the voltage & current measuring "millivolt & micro amps" value is accuracy. From the 1000uf and 1.5F capacitors. Voltage sensor: The two resistors are R1 100K and R2 10K in the range (0-30V).

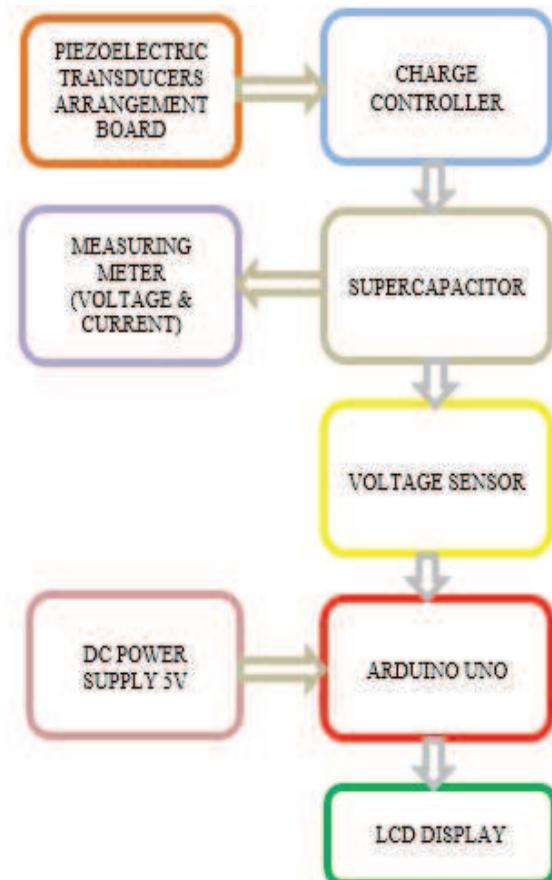


Fig.1. Block diagram

Arduino Uno: The CPU (focal preparing unit) of our project is a microcontroller. We're using an ARDUINO UNO in this case. The various components of a microcontroller are similar, □ Reading voltage from Super capacitor.

- Sending this perusing voltage information to LCD show. Power Supply: 9V supply for ARDUINO 5V & 16x2 LCD display 5V.

III. WORKING MODEL & STRUCTURAL SPECIFICATIONS

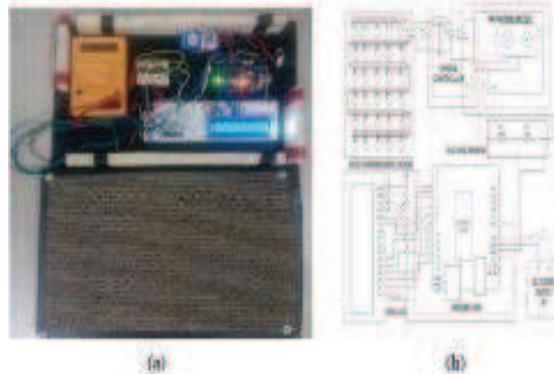


Fig.2. (a).hardware setup (b).key diagram

When we walk or jump, we generate vibrations that are transferred to the charging board (piezo) beneath the crosswalk, which collects energy from these vibrations. This technology converts kinetic energy into electrical energy, which is then stored and used for various purposes; the energy generated by the piezoelectric panel can charge a super capacitor (which can be used further). We use the Charge controller's efficient output energy, which is stored in the super capacitor. We use a voltage divider (voltage sensor) to measure the super capacitor voltage value. The sensor sends the analogue signal to Arduino, which has a microcontroller and other electronic components that convert the analogue signal to a digital signal, which is then connected to an LCD display.

it is the monitoring circuit measuring voltage value message is display in LCD display. This product is best used in densely populated areas. For densely populated countries, the use of excess energy from human base power is critical. Where human velocity is high, such as in crowded metro stations, train stations, transportation hubs, airports, and sanctuaries. Piezoelectric power generation was installed on the sidewalks, stairwells, highways (vibration from moving vehicles), athletes' running ground, jogging ground, and gym& fitness centre. The scholars' test case was shown in real time at the Venice Merino and two train stations in Torino, Italy. Lead is used as a piezoelectric transducer material in this paper the piezoelectric transducer made of zirconate titanate (PZT) is used to convert tension or mechanical pressure into an electrical power substitute. The piezoelectric transducer uses a piezoelectric material that has an unusual property, such as the material prompting voltage when tension or stress is applied to it, and it is used to estimate actual amounts such as power, pressure, and stress, among other things. Depending on our required flow or voltage

esteems, a piezoelectric transducer may be used in a series equal blend in the activity. One piezoelectric transducer generates 10-20 mV of voltage and 100 uA of current. Piezoelectric transducers create ac, and we use a span rectifier (IN4148) to convert it to dc. These diodes are small sign diodes with high exchanging activity, and they convert the piezo voltage to dc, which is then used to charge the super capacitor. Extension rectifiers do not produce pure dc (a few waves appear). To obtain pure dc, we employ a capacitor (1uf, 63V), which eliminates the waves and produces pure DC. To acquire the best yield, a bridge rectifier and capacitor are used with each piezo transducer in this project.



Figure 3 Model of Piezoelectric transducers arrangement board

IV. PRESSING MECHANISM & CHARGE

CONTROLLER

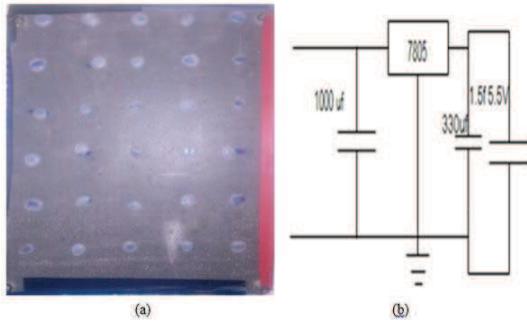


Figure 4 fixed position and control circuit

Rubber mat ($L=30\text{cm}$ * $B=25\text{cm}$) is used for footstep pressing under the mat hot glue point for squeeze the transducers point pressure, touch the piezoelectric transducers electrode. By this process of pressing mechanism piezoelectric transducers cannot damage at high pressure and advantage is long durable. The Piezoelectric transducer power age interface circuit is the Charge regulator and its schematic graph displayed in Figure.3. The comparing block outline is displayed in Figure.4.

The improvement of Charge regulator gives a circuit component to get a legitimate yield while

input voltage is low to store the super capacitor. This circuit is very efficient.

V. CALCULATION

We measure the voltage across the capacitor in our piezoelectric generators for two reasons. The first is just to ensure it is working. The second, more significant explanation, is on the grounds that we can utilize this voltage to compute the measure of energy put away in our capacitors utilizing the

$$\text{Condition: } E=1/2 V^2 C \quad (1)$$

Where E is the energy put away in the capacitor, C is the capacitance and V is the voltage estimated across the capacitor. When utilizing this condition, ensure the units are right. We need the capacitance to be μF we need the voltage to be in volts (V). Then, at that point, the units of energy will be joules (J) or proportionally, in watt-seconds (w/s). To work out how much energy is put away for each tap or push on the piezo component, measure the voltage previously (V_0) and the voltage after the tap (V_1), and afterward utilize the accompanying condition:

$$E_{\text{Step}}=1/2([V_1]^2 - [V_0]^2) C J V_1, \dots, V_\infty, \text{ (micro joules)} \quad (2)$$

$$[E_{\text{Step}}=\text{Micro joules to joules}] 1 \text{ micro joules} = 1 \times 10^{-6} \text{ joule} \quad (3)$$

1000uf Capacitor Calculation. In 1000uf capacitor piezoelectric transducers average 80kg person per step charge stored value is 30 mV and 30 to 40 μA . From eqn (2)

$$[E_{\text{Step}}=1/2([V_1]^2 - [V_0]^2) C J V_1, \dots, V_\infty, [E_{\text{Step}}=1/2([0.06]^2 - [0.03]^2)]$$

$$(1000\mu\text{f})] V_1, \dots, V_\infty,$$

$$E_{\text{Step}} = 1.35 \text{ micro joules}$$

From eqn (3)

$$[E_{\text{Step}}=\text{Micro joules to joules}] 1 \text{ micro joules} = 1 \times 10^{-6} \text{ joule}$$

$$[E_{\text{Step}}=(1.35) \text{ Micro joules to joules}] 1 \text{ micro joules} = 1 \times 10^{-6} \text{ joule}$$

$$E_{\text{Step}} = 1.35e-6 = 1.35 \times 10^{-6} = 0.00000135 \text{ joules/step}$$

5.3.2 1.5F 5.5V Super Capacitor Calculation

In 1.5F 5.5V Super capacitor charge stored value is 0.02V (V_1) and 0.01V (V_0) 10mV can charge time taken 5minutes. From (2)

$$[E_{\text{Step}}=1/2([V_1]^2 - [V_0]^2) C J V_1, \dots, V_\infty, \text{ Convert 1.5F to } 1500000 \mu\text{f}]$$

$$[E_{\text{Step}}=1/2([0.02]^2 - [0.01]^2) (1500000)] V_1, \dots, V_\infty,$$

$$E_{\text{Step}} = 225 \text{ micro joules}$$

From (3)

$$[E_{\text{Step}}=\text{Micro joules to joules}] 1 \text{ micro joules} = 1 \times 10^{-6} \text{ joule}$$

$$[E_{\text{Step}}=(225) \text{ Micro joules to joules}] 1 \text{ micro joules} = 1 \times 10^{-6} \text{ joule}$$

Formula- micro joules x 10-6 joule

Estep =0.000225 joules/step

As we probably are aware the tension is straightforwardly corresponding to measure of force created

Po Wt. At this time we approve to take the constant of proportionality as K, then, at that point, the condition becomes

We realize that for wt. =80Kg, we get the worth of voltage $V=1.30v$ and $I=0.0002A$ Then, at that point, $P=V*I= 1.30*0.0002 = 0.00026w$, implies we can approximately that for 80kg we get power (P) =0.26 mw. As of this we can determine the worth of K , $K=P/wt. = 0.00026/80 K = 0.00000325$, transducers cannot damage at high pressure and advantage is long durable. The Piezoelectric transducer power age interface circuit is the Charge regulator and its schematic graph displayed in Figure.4.10 The comparing block outline is displayed in Figure 3.1. The improvement of Charge regulator gives a circuit component to get a legitimate yield while input voltage is low to store the super capacitor. This circuit is very efficient.

VI. RESULTS

Graph For Per Step {1000 μ F Capacitor}
VOLTAGE V

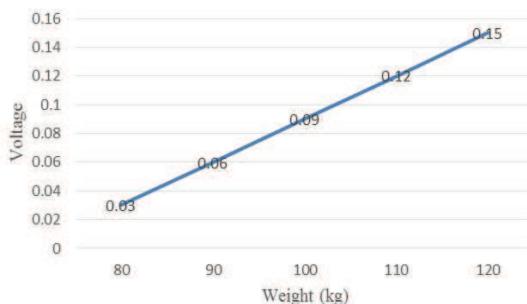


Figure 5 Voltage

MILLIAMPS mA

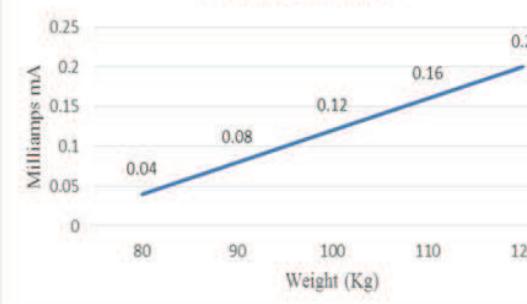


Figure 6 Milliamps

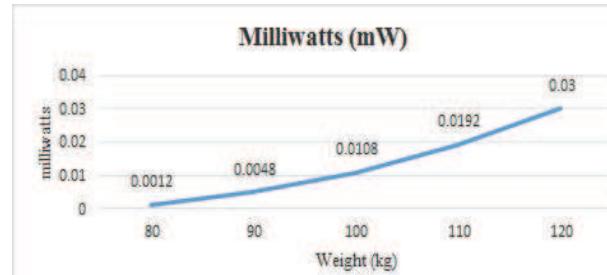


Figure 7 Milliwatts

VII. CONCLUSION

The power generation of piezoelectric transducers for low-power applications has been successfully evaluated and implemented, making it the most cost-effective and accessible energy alternative for the general public. Renewable energy does, in reality, add to our primary energy. If this research is implemented, we will not only be able to solve the energy crisis, but we will also be contributing to a healthy worldwide environmental transformation. It's a simple mechanism.

REFERENCES

- [1]. Priyanka Baniya1, Shivanjali Barge2, Rahul Patil3, Prof. Poonam Yewale4 May 2019 "Power Generation using Piezoelectric Material." International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 05.
- [2]. Muhammad Aamir Aman, Hamza Umar Afridi, Muhammad Zulqarnain Abbasi, Akhtar Khan, Muhammad Salman September-October (2018) "Power Generation from Piezoelectric Footstep Technique." Mech Cont.&Math. Sci., Vol.-13, No.-4.
- [3]. Aswal, Pankaj & Singh, Suyash & Thakur, Apurv. (2017) "Generation of Electricity by Piezoelectric Crystal in Dance Floor." 10.1007/978-981-10-1708-7_52.
- [4]. Jaskaran Kaur, Abhinav Vishnoia May -2016 "Review on Piezoelectric Smart Road Sensor." International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-4, Issue-5.
- [5]. Frontoni, Emanuele & Mancini, Adriano & Zingaretti, Primo & Gatto, Andrea. (2013) "Energy Harvesting For Smart Shoes: A Real Life Application." 4.10.1115/DETC2013-12310.
- [6]. Henry A. Sodano, Daniel J. Inman, Gyuhae Park May 17, 2004 "Estimation of Electric Charge Output for Piezoelectric Energy Harvesting." LA-UR-04-2449, Strain Journal, 40(2), 49-5S, 20.

Coronary Artery Disease Prediction using Machine Learning Algorithms

Dr Baranidharan Balakrishnan
SRM Institute of Science and Technology,
Chennai, India
baranidb@srmist.edu.in

Adhith Sankar
SRM Institute of Science and Technology,
Chennai, India
as2389@srmist.edu.in

Kaushik Kumaran
SRM Institute of Science and Technology,
Chennai, India
kk1234@srmist.edu.in

Keshav Rao
SRM Institute of Science and Technology,
Chennai, India
kr7559@srmist.edu.in

ABSTRACT - In the past decade, Cardiovascular Disease (CVD) has been one of the major death factors for humanity. Early diagnosis could increase the chances of curing patients of their disease and will help in increasing their impermanence. Contemporarily, Machine Learning algorithms are being widely used in the medical community and have been giving assuring results which has had an impact on saving lives as well as saving time. To attain better accuracy, Machine learning models are specifically trained using various datasets. Through this research, our primary intention is to obtain higher accuracies using various machine learning models that have been developed by training these datasets pertaining to heart disease prediction based on various metrics. We intend to reduce the financial cost and time by avoiding costly medical examinations. This reduces the waiting time it takes for the patients to undergo multiple such tests. In this paper, it is demonstrated that CVD can be predicted using simple medical tests, supported with advanced machine learning algorithms.

Keywords: CVD, Machine learning, Random Forest, Decision Tree.

I. INTRODUCTION:

Cardiovascular Disease (CVD), is considered to be one of the deadliest diseases we have come across as it leads to cardiac arrests [1]. Due to the rare occurrence and symptoms that can be often related to other diseases, many people do not take the risk of heart failure seriously. There are different ways by

which the heart might fail, such as due to damage to the heart arteries and a sudden sensation of a squeeze in the heart can be felt. Ahmad and G.Wang [2] used CVD as a collective term for any form of heart disease.

Any condition that affects the heart and its vessels can be related to coronary vascular diseases. CVD is more frequently found amongst men than women. Considering the contemporary rate at which cardiovascular disease is increasing, it imposes a huge financial burden.

There are many ways to identify heart diseases. The traditional way to do this is by identifying them through ECG or Electrocardiogram tests and can also be identified by using Angiogram. The deformities in the human heart can be easily identified using ECG by medical experts Acharya U et al [3].

Artificial Intelligence plays a vital role in giving a diagnosis and smart decision making. These diagnoses can be categorized using various classifiers. The Health industry generates copious amounts of medical data which contains a lot of beneficiary information. Javad Hassannataj Joloudari [4] and Raghupathi [5] analyzed and experimented on many classifier models and achieved predictive diagnosis avoiding unnecessary cost for the medical sector.

Due to its unaffordability, it is not easily accessible for economically weaker sections of society. To overcome such a circumstance and to make

expensive healthcare widely available and affordable to all people, a quintessential model has to be developed.

We have used these popular classifiers and they are, Logistic Regression, K-Nearest Neighbor, Support Vector, Naive Bayes, Decision Tree, Random Forest, MLP Classifier, XG Boost and Gradient Boost.

Given below are few technical explanations of the various classifiers:

- **Logistic Regression:** Probability is the main concept being implemented here. It provides value in binary format, i.e. 1 and 0.
- **K Nearest Neighbor:** KNN is a non-parametric algorithm that makes no assumptions about the input data. It collects the data which are available and classifies new data points based on how similar they are.
- **Support Vector Machine:** The way SVM works is that it maps data into a multidimensional feature space, allowing data points to be categorized even when the data cannot be separated linearly.
- **Naive Bayes:** It is based on the Bayes Theorem and conditional probability. This classifier is based on the assumption that it is independent between attributes of data points.
- **Decision Tree:** This classifier is used for its capability to capture descriptive decision making knowledge from the supplied data. In layman terms, the model asks a question and based on the answer which will be either ‘Yes’ or ‘No’, it further splits the tree into subtrees.
- **XGBoost:** It uses the same machine learning techniques as Gradient Boosting, this classifier also has a parallel tree boosting that is used for enhancing accuracy.

- **Random Forest:** This classifier consists of a sizable number of individual decision trees. Each tree is being diagnosed and the set of trees with the most votes is elected as the model’s prediction.

- **Multi-Layer Perceptron:** MLP is a feedforward AN model which takes an input data set to a corresponding output dataset. It consists of multiple layers, each fully connected to the next one.

- **Gradient Boost:** In this boosting classifier, each predictor tries to improve from its previous results by reducing the errors.

II. LITERATURE SURVEY:

There are various Machine Learning techniques that are used to determine and analyze heart diseases. Alizadehsani et al [6] have built various classification models which use Bagging, Artificial Neural Networks and Naives Bayes to detect Coronary Artery Diseases.

Javad Hassannataj Joloudari [4] had found that Database Knowledge Discovery (KDD) is the ideal method as the highest rates of disease diagnosis can be obtained in the current given situation of the healthcare sector. The drawback in using the KDD model is the feature selection model, that requires us to select the optimal subset from the dataset using which the model was trained which is difficult as the subset should not affect the accuracy and not many must be removed as that would alter the results.

L Parthiban [7] had found that coactive Neuro-fuzzy modeling was a solid and dependable method for the mapping between the different attributes. HRV models were observed by Melillo et al [8] to determine CAD and concluded that it performed way better than Echographic parameters.

Sellappan Palaniappan [9] proposed that an IHDPs (Intelligent Heart Disease Prediction System) can be developed using Naive Bayes, Neural Network and

Decision Trees. Srinivasan [10] had used a Decision Tree as it had provided better results when compared to other classifiers and this led to the development of a heart care system.

Baranidharan B [11] have done the comparative research metrics of various classifiers on Cleveland and Statlog dataset and have found out that Naive Bayes classifier gives the best accuracy of 84% without using any bagging or boosting techniques.

Based on the above survey, we have come to the conclusion that the MLP, XGBoost and Gradient Boosting classifiers are the ideal approaches.

Gradient Boosting is a type of classifier which works by trial and error method. It learns and improvises from its previous errors are used together with many weak learning models, to create a stronger diagnostic model. XGBoost is a more structured form of Gradient Boosting. Its training is very fast and can be parallelized across clusters.

Table 1 shows the different accuracies, Precision, Recall, F1 Score we were able to obtain using different models. In this paper, we have used the datasets of **Cleveland, Hungary, Switzerland, and Long Beach**, which when combined have 1000+ data points. It provides us with data such as age, sex, the type of chest pain, resting blood pressure, cholesterol, fasting blood sugar, number of major vessels colored by fluoroscopy, thalassemia, diagnosis of heart disease and much more.

III. EXPERIMENTAL SETUP:

The classifiers made use of in this research project have taken the help of hardware resources with configuration of Intel i7 Processor 8th generation, 4 GB RAM, Windows 10 using the Jupyter Notebook Environment. Classifiers which are inbuilt from SciKit library are utilized in this research. Many AI specialists have made use of the UCI Coronary illness dataset [12] which is gathered from **Cleveland, Hungary, Switzerland, and Long Beach**. Different types of classifier models are observed during this research.

Table 1. Comparison of various classifiers:

S. No	Classifiers	Accuracy	Precision	Recall	F1 score
1.	Logistic Regression	0.8390	0.9142	0.8	0.8533
2.	K- Nearest Neighbors (k=2)	0.9285	0.8797	0.9788	0.9266
3.	Support Vector ('linear')	0.8390	0.9238	0.7950	0.8546
4.	Support Vector ('rbf')	0.7170	0.7809	0.7008	0.7387
5.	Multi-Layer Perceptron	0.8585	0.9238	0.8333	0.8675
6.	Naïve Bayes	0.8292	0.8761	0.8070	0.8401
7.	XGBoost	1.0	1.0	1.0	1.0
8.	Decision Tree	1.0	1.0	1.0	1.0
9.	Random Forest	1.0	1.0	1.0	1.0
10.	Gradient Boosting	0.9756	0.9809	0.9716	0.9763

The results of all the classifier models are presented in Table 1 (given above) where you can see that Gradient Boosting Classifier gives the best accuracy of all the classifier models which is 97% followed by K – Nearest Neighbors ($k = 2$) which gives 92%. These two models are considered as the best performing models among all classifiers (without using bagging or boosting techniques). Support Vector (linear kernel) is the only other classifier with an accuracy over 90%. Coincidentally the least accurate classifier is the Support with a ‘rbf’ kernel.

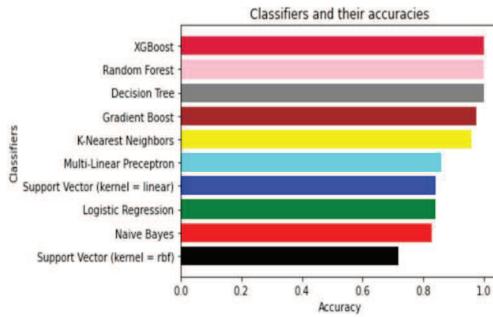


Figure 1: Accuracies of classifiers using 13 attributes.

In Table 1, it is found in the Support Vector Machine algorithm, the linear kernel performs better than the rbf kernel. This is because the data is linearly separable between the hyper-plane’s regions. We can also come to the conclusion that XGBoost, Random Forest and Decision Tree give 100% accuracy. This is because all three algorithms are specifically designed to give better results and especially when the number of data is considerably small which is 1025, it is only more likely that it gives 100% accuracy.

Using the above data, we were able to determine the classifiers which gave the best results. Using those classifiers, we determined which data columns contribute the most for the accuracy and we found out that CP (Chest Pain), CA (Number of major vessels) and thal (Thalassemia) contributes the most which is given in the below diagram.

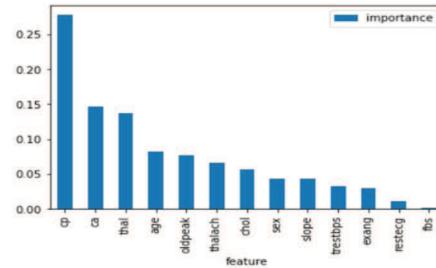


Figure 2: Ranks of the attributes based on Accuracy.

By referring to the above data, the attributes CP (Chest Pain), CA (Number of major vessels), thal (Thalassemia), age, Old Peak, and Thalach were all used to create a new dataset. Baranidharan B [13] had used the same technique to find the most-contributing attributes using the Cleveland and Statlog datasets and after comparing the results it is inferred that two attributes (age and thalach) are different. This is because the dataset which we are using has more/additional data points which influences these factors.

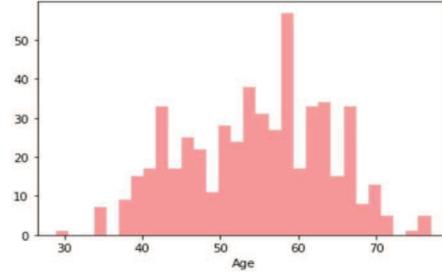


Figure. 3: Patient's age vs Number of Positive Diagnosis

This was done in order to reduce the number of parameters which would in turn require the patients to take fewer tests that do not have an affiliation to the required result and all its data points were taken into consideration and the model was trained and its results are given as below.

Table 2. Comparison of various classifier models on the new dataset:

S. No	Classifiers	Accuracy	Precision	Recall	F1 Score
1.	Logistic Regression	0.8181	0.8924	0.7833	0.8343
2.	K- Nearest Neighbors	0.9123	0.8481	0.9781	0.9084
3.	Support Vector (kernel ‘linear’)	0.8116	0.8924	0.7747	0.8294
4.	Support Vector (kernel ‘rbf’)	0.7207	0.7468	0.7195	0.7329
5.	Multi-layer Perceptron	0.8376	0.8987	0.8068	0.8502
6.	Naïve Bayes	0.8506	0.8924	0.8294	0.8597
7.	XGBoost	0.9902	0.9810	1.0	0.9904
8.	Decision Tree	1.0	1.0	1.0	1.0
9.	Random Forest	1.0	1.0	1.0	1.0
10.	Gradient Boosting	0.9415	0.9430	0.9430	0.95

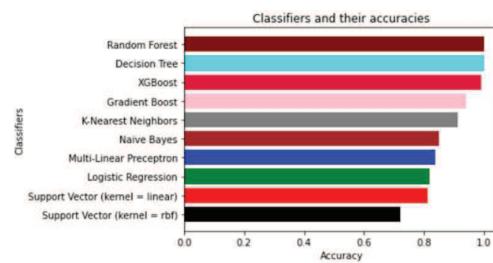


Figure 4: Accuracies of various classifiers using new dataset.

After checking the results, it is inferred that just these 6 attributes are enough to predict heart disease by reducing the number of tests/scans required to get a proper diagnosis and this makes it affordable for all the sections of the society. Baranidharan B [13], had identified 6 best attributes based on Extra Tree Classifier. When mapped and compared the classifiers between Table 1 and Table 2, it is noticed that their respective accuracies vary very little and some classifiers like Naive Bayes perform even better with this new dataset.

IV. EXPERIMENT RESULTS:

Concluding the results from Table 1, we can observe that XGBoost, Decision Tree and Random Forest give 100% accuracy without getting introduced to any Bagging and Boosting algorithms. Following that both Gradient Boosting and K-Nearest Neighbors give 97% and 92% accuracies.

In Table 2, we have used the new dataset with 6 attributes that contributes mostly to the accuracy, we can observe that the results are comparatively minimal when compared with Table 2. Both Decision Tree and Random Forest give 100% accuracy but XGBoost reduces by 1%. Even Though, the accuracies of some classifiers are reduced by 1% or

2%, classifiers such as Gradient Boosting and Multi-Layer Perceptron show better results with an increase in accuracy by 3% and 2% respectively.

We can conclude from the above results; we obtain the best accuracy from both the Decision Tree model and Random Forest model compared to the other models. This is because it uses the most important attributes as the major factor to give accurate results. Random Forest gives the same result since it is a derivative of Decision Tree.

V. CONCLUSION AND FUTURE WORK:

In this proposed work, we have been able to achieve higher levels of accuracy using classifiers such as Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Naïve Bayes. The fundamental basis of this research are: (i) To compare the performance of different classifiers using the heart disease dataset collected from four different databases: Cleveland, Hungary, Switzerland and Long Beach, and (ii) To reduce the cost and make it affordable and minimize the number of tests/scans the patient has to undergo.

In future implementations, heart disease can be classified and prioritized according to their risk levels and can be experimented with different parameters.

REFERENCES

- [1] World Health Organization. *Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth*. World Health Organization, 2017.
- [2] Ahmad, Gulzar, Muhammad Adnan Khan, Sagheer Abbas, Atifa Athar, Bilal Shoaib Khan, and Muhammad Shoukat Aslam. "Automated diagnosis of hepatitis b using multilayer mamdani fuzzy inference system." *Journal of healthcare engineering* 2019 (2019).
- [3] Acharya, U. Rajendra, Oliver Faust, Vinitha Sree, G. Swapna, Roshan Joy Martis, Nahrizul Adib Kadri, and Jasjit S. Suri. "Linear and nonlinear analysis of normal and CAD-affected heart rate signals." *Computer methods and programs in biomedicine* 113, no. 1 (2014): 55-68.
- [4] Joloudari, Javad Hassannataj, Edris Hassannataj Joloudari, Hamid Saadatfar, Mohammad Ghasemigol, Seyyed Mohammad Razavi, Amir Mosavi, Narjes Nabipour, Shahaboddin Shamshirband, and Laszlo Nadai. "Coronary artery disease diagnosis: ranking the significant features using a random trees model." *International journal of environmental research and public health* 17, no. 3 (2020): 731.
- [5] Raghupathi, Wullianallur. "Data mining in healthcare." *Healthcare Informatics: Improving Efficiency through Technology, Analytics, and Management* (2016): 353-372.
- [6] Alizadehsani, Roohallah, Jafar Habibi, Mohammad Javad Hosseini, Hoda Mashayekhi, Reihaneh Boghrati, Asma Ghandeharioun, Behdad Bahadorian, and Zahra Alizadeh Sani. "A data mining approach for diagnosis of coronary artery disease." *Computer methods and programs in biomedicine* 111, no. 1 (2013): 52-61.
- [7] Parthiban, Latha, and R. Subramanian. "Intelligent heart disease prediction system using CANFIS and genetic algorithm." *International Journal of Biological, Biomedical and Medical Sciences* 3, no. 3 (2008).
- [8] Melillo, Paolo, Raffaele Izzo, Ada Orrico, Paolo Scala, Marcella Attanasio, Marco Mirra, Nicola De Luca, and Leandro Pecchia. "Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis." *PloS one* 10, no. 3 (2015): e0118504.
- [9] Palaniappan, Sellappan, and Rafiah Awang. "Intelligent heart disease prediction system using data mining techniques." In *2008 IEEE/ACS international conference on computer systems and applications*, pp. 108-115. IEEE, 2008.
- [10] Srinivas, K., G. Raghevendra Rao, and A. Govardhan. "Analysis of coronary heart disease and prediction of heart attack in coal mining regions using data mining techniques." In *2010 5th International Conference on Computer Science & Education*, pp. 1344-1349. IEEE, 2010.
- [11] Balakrishnan, Baranidharan. "A Comprehensive Performance Analysis of Various Classifier Models for Coronary Artery Disease Prediction." *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)* 15, no. 4 (2021): 1-14.
- [12] Blake, C., and C. J. Merz. "UCI repository of machine learning databases, Department of Information and Computer Science, University of California, Irvine, CA, 1998." URL: <<http://www.archive.ics.uci.edu/ml> (2015).
- [13] Baranidharan, B., Abhisikta Pal, and Preethi Muruganandam. "Cardiovascular disease prediction based on ensemble technique enhanced using extra tree classifier for feature selection." *International Journal of Recent Technology and Engineering (IJRTE)* 8, no. 3 (2019): 3236-3242.

Estimation of Soil Salination for Crop -Yielding and How to Prevent

¹Reshma Neppalli, ¹Ch.Gowthami Supriya, ¹A.Hema, ¹K.Kiran Kumar, ²N.V.K Ramesh and ^{*3}B. Naresh Kumar Reddy

¹Department of Computer Science and Engineering, KL.University, India.

²Department of Electronics and Communication Engineering, KL.University, India.

^{*3}Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India.

Email: naresh.nitg@gmail.com,

Abstract—Now-a-days the farmers are facing various crop yielding problems due to climatic conditions and lack of fertility in soil salinity has negative impacts in growth of plants and causes destruction of land. Because of less salinity, the yield of agricultural products is lower as it is not favorable for farmers. It even destroys the economic growth of the region. Detecting the soil salination at an early stage helps in getting good production. The complete loss of farmlands is due to heavy contamination, and it has negative effects of soil erosion. Moreover, some crops are not suitable in certain areas due to increase of salt content in water. This method is proposed to identify the salination in the soil and suggest suitable crops based on the salination level. In-order to increase the farmer's academic growth, this method is proposed. In this project we collected the pH values of different soil and suggest the suitable crops for the soil based on pH value. Agriculture depends on the soil and based upon the soil we can estimate whether the crop is suitable for the cultivation or not based on the soil. The soil comprises of minerals like S, Cl, Ph, NO₃. The composition of the correct of these minerals leads to the good yield of the crop. So, it is important to check the soil and estimate the mineral content accordingly so that we can estimate the crop based on the mineral content. Salination is nothing but the collection of exact number of soluble salts in the water. Plant growth gets effected by soil salinity due to the meddle of the salt in water absorption. That is the reason even when sufficient soil moisture was provided also the crops die as they were unable to take enough water. Therefore, because of this soil salination, many lands were being excluded for the agriculture. Hence, using the pH sensor the solution was proposed to this problem.

Index Terms—NodeMCU . Arduino IDE (Integrated Development Environment), pH Sensor, Soil Salination, USB.

I. INTRODUCTION

Agriculture plays a leading role in our economic life. It is considered as the backbone of the economic system of our country. Basically, Agriculture depends on the soil [1]. Based on the soil texture and the fertility, we can estimate the crop which has to be cultivated or whether the land is suitable for cultivation or not. The soil comprises of different minerals like primary minerals and secondary minerals. Primary minerals are the minerals which cannot be chemically metamorphosed since the decay. Silicate, apatite, volcanic gases etc. are the primary minerals [2]. Secondary minerals are the minerals found in the soil which were formed with the disintegrating of primary soil mineral like Carbonates (CO₃), S(Sulphur),

Cl(chlorine), Ph (Phosphorous), NO₃(nitrate), Halides, hydroxides etc. are the examples of secondary minerals [3]. So, one of the reasons why the soil cannot cultivate the crops is the content of NaCl(sodium) level is high. It is also known as salination [4].

Salination is the lodgment of the salts soluble in water. Soil salination can be formed with the processes like evaporation, dissolution, precipitation, salt transport, ion exchange etc. like other processes [5]. The soils which are affected by salinity contains immoderate concentrations of either soluble salts or exchangeable Na or it can be with which leads to the deficient percolating of base forming cations [6]. So, the most common salt affected areas are: Saline –sodic, saline, sodic. Due to the interfere of the salt in the water absorption, the plant growth gets effected by soil salinity which is the reason for destruction of crops as they unable to take enough water even they have sufficient soil moisture. Soil salination can be caused because of the retaliation of the soluble salts in the earth surface and this may be happening in either natural way or due to the inappropriate organic activities like farming practices, low salt dissolution and removal.

The causes for soil salination:

- The soil salination occurs when the extreme salts were not flushed from the earth surface properly and thus dry climate and low precipitation occurs.
- Due to the high evaporation rate, the salt was collected at the ground surface.
- The excess water in the root zone which was accompanied by anaerobic conditions and the water system with poor drainage system will occur when the salts were not flushed or washed properly due to the lack of the transportation for water.
- The increase of salt level in the earth surface is due to the saltwater irrigation.
- The salt level can be increased in the soil or earth surface due to the removal of the raised water table and deep – rooted vegetation.
- The salt salination occurs when the amount of salt gets penetrated the ground water and in the leakages which form from the geological deposits.

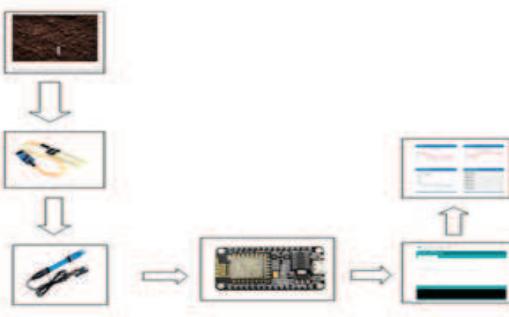


Fig. 1. Block Representation.



Fig. 2. Hardware Connection.

- The salt level gets raised when the salt gets seeped into the lowlands.
- Gusts from the coastal areas will blow salty air mass to coastal territories nearby which causes the salt salination.
- Sea water overflow which was occurred due to the salt evaporation causes the salt salination in the soil surface.
- Due to the improper usage of the pesticides and fertilizers, the chemicals get percolated in the soil surface and gets converted into the salts which can be either soluble or insoluble. This leads to the salt salination of the soil.

II. LITERATURE REVIEW

A semi analytical model using an exponent function was developed to estimate soil salt content under different moisture conditions based on the control laboratory experiment [7]. The second phase is the root mean square and the mean square which were given the values were taken and the model which was taken is used to estimate the soil salt content in the wet soil [8]. The next phase is that considering the effects of both soil moisture and soil salt on soil reflectance sum up will give the output which is the semi analytical model which was soil salt content estimated error. The approach presented in this paper provides a new way of estimating soil salinity from soil spectra under various soil moisture conditions, and it will be a potential application for large -scale SSC (Sector Skills Council) mapping [9].

III. PROPOSED MODEL

The block diagram shown describes the methodology for this proposal, shown in Fig. 1. First, we use the soil sensor to check the moisture and other factors in the soil so that we can find the salt salination in the soil according to those related factors. Then we use the PH sensor to find the salt pH content in the soil. These sensors were connected to the open-source firmware which is considered as Nodemcu. It runs ESP8266 Wi-Fi-SoC and esp-12 hardware. Nodemcu means node associated microcontroller .it undergoes the prototyping board designs which is used in the circuit board functioning as a dual in-line package which integrates a USB controller with MCU, antenna. This nodemcu was associated with the

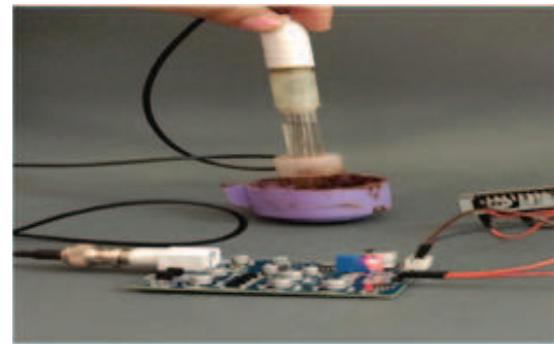


Fig. 3. Hardware Connection1.

Arduino IDE (Integrated Development Environment) which is an Arduino software that contains an editor which is used to write code. It contains text console and a toolbar with buttons which is used for accessing the functions. It even contains Arduino hardware which is used to upload programs and helpful to communicate with them [2] into the cloud platform .Thing Speak software is an open – source software which allows users to communicate with the devices [3].Thing Speak provides the visualizations of the data that has been posted by our devices to the platform [4].So the Arduino IDE (software) acts as between the nodemcu(hardware), and thing Speak(cloud) and stores the data into the cloud. According to the data that has been measured with the sensors in the soil, the salt pH of the soil has been regulated for distinct kinds of soils. Hence according to those results, the crops have been suggested so that according to that moisture content and other conditions, the crops have been suggested and hence can be cultivated in that soil. hence this is our solution to use lands that has been affected by the salinity instead of being excluded.

IV. HARDWARE IMPLEMENTATION

For this project, we divided the hardware and software implementation as the hardware consists of collecting the data from the nodemcu, pH sensor and the several types of soil which is a hardware input [10]-[25]. So, for the estimation of

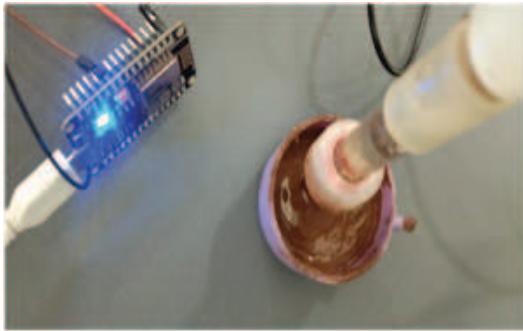


Fig. 4. Hardware Connection2.

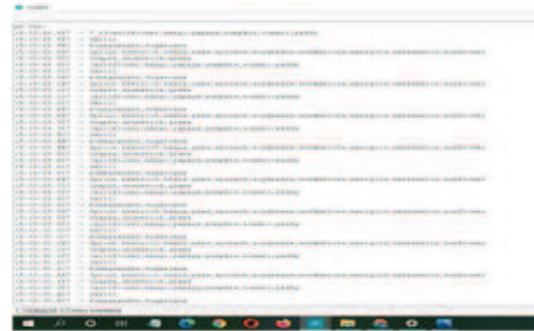


Fig. 6. Crops estimated for soil sample 1.

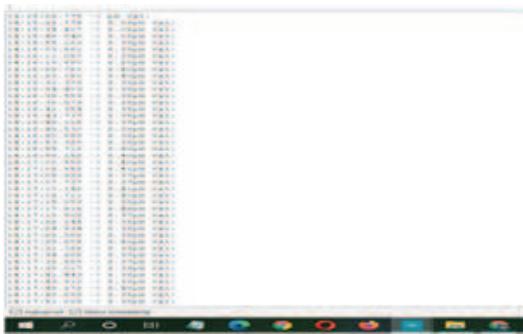


Fig. 5. pH values of soil sample 1 output.

soil salination, the NodeMcu related to the pH sensor using the jumper wires and for the Ph sensor we need the battery of 3V, and a potentiometer was even included when the values get deviated from the path due to the overloading of the data shown in Fig. 2, Fig. 3. and Fig. 4. Then with this setup we calculate Ph of the salt in different soils like red soil, black soil, and other soils from various places. In the Arduino software, the respect libraries for the nodemcu like ESP8266 etc., for Thing Speak (cloud Platform) were being included. After that C language code was being developed for calculating the pH of salt in different soils and updating them in the Thing Speak cloud. This is actual the problem statement for our project, and we suggested the solution of prescribing the crops according to the pH content which is matching to the pH content of the soil. We will develop the code for prescribing the crops accordingly and that needs to get updated in the cloud so that it would be helpful for the farmers to cultivate the crops instead of the leaving the land.

V. EXPERIMENTAL RESULTS

The Proposed system is developed, and the developed system is tested with several types of soil samples which are collected from the surrounding areas. We have calculated the value of pH of various soils shown in Fig. 5. By getting the value of pH from soil then we suggest the suitable crop for the

No	Crops	Optimum pH	Min pH	Crops	Optimum pH
1	Banana	6.0-7.0	5.0-6.0	Potato	4.5-5.5
2	Bitter gourd	6.0-7.0	5.0-6.0	Prunes	5.0-6.0
3	Black gram	6.0-7.0	5.0-6.0	Rhubarb	5.0-7.0
4	Bitter melon	6.0-7.0	5.0-6.0	Raspberries	5.0-7.0
5	Carrot	6.0-7.0	5.0-6.0	Sorghum	5.0-7.0
6	Cauliflower	6.0-7.0	5.0-6.0	Soybean	5.0-7.0
7	Cowpeas	6.0-7.0	5.0-6.0	Corn	5.0-7.0
8	Cowpea/Cowpea	6.0-7.0	5.0-6.0	Okra	5.0-6.0
9	Cowpea/Oxpea	6.0-7.0	5.0-6.0	Orchard grass	5.0-6.0
10	Cowpeas	6.0-7.0	5.0-6.0	Phenophenol	5.0-6.0
11	Cowpea	6.0-7.0	5.0-6.0	Wheat	5.0-6.0
12	Cotton	6.0-7.0	5.0-6.0	Wild rice	5.0-6.0
13	Cabbage	6.0-7.0	5.0-6.0	Yield	5.0-6.0
14	Flax	6.0-7.0	5.0-6.0		
15	Garbanzo	6.0-7.0	5.0-6.0		
16	Green beans	6.0-7.0	5.0-6.0		
17	Green pea	6.0-7.0	5.0-6.0		
18	Mustard	6.0-7.0	5.0-6.0		
19	Mango	6.0-7.0	5.0-6.0		
20	Pistachio	6.0-7.0	5.0-6.0		

Fig. 7. Crops estimated according to pH of the soil salinity.

soil. Different values of pH have various crops. When the pH is displayed on the Arduino IDE then we will get the suggested crop following the displayed pH. The entire pH values are uploaded to the Think Speak.

First, we have tested with red soil to get the pH values then we can suggest the suitable crop for the soil shown in Fig. 6. When we have calculated the pH value, the entire data is collected, and values are uploaded into the think Speak. We have used different soils for testing they are red soil, black soil, sand, alluvial soil, and clayey soil shown in Fig. 7.



Fig. 8. Graph for the soil sample.

The outputs that are displayed on the arguing IDE and hence we got the Ph values. The suitable crops also displayed along with the pH. So, it helps in proper use of soils. The pH values were measured according to the pH values list of the crops shown in Fig. 8.

VI. CONCLUSION

An economically affordable soil testing was done by using Arduino IDE for showing the values of pH, nodeMCU for transferring of data from sensor to IDE and to think Speak and pH sensor for getting values. This is very much helpful for the crops where the soil contains the high salinity level like the crops near the seashores.

REFERENCES

- [1] Julián Cuevas, Ioannis N. Daliakopoulos, Fernando del Moral, Juan J. Hueso and Ioannis K. Tsanis, A Review of Soil-Improving Cropping Systems for Soil Salinization, *Agronomy*, Vol 295, 2019.
- [2] B.Naresh Kumar Reddy, B.Siva Hari Prasad, B.Y.V.N.R. Swamy,Innovative Water Saving Agriculture by Using Resources” IJECET, Pp227-237, Vol.3, Issue 2, July-Sept 2012.
- [3] B. Naresh Kumar Reddy, K Sarangam, T Veeraiah, R Cheruku , SRAM cell with better read and Write stability with Minimum area, *IEEE Region 10 Conference (TENCON)*, 2164-2167, 2019.
- [4] Shakeel Ahmed, N. V. K. Ramesh and B. Naresh Kumar Reddy, A Highly Secured QoS Aware Routing Algorithm for Software Defined Vehicle Ad-Hoc Networks Using Optimal Trust Management Scheme, *Wireless Personal Communications*, 2020
- [5] B. Naresh Kumar Reddy, B. Veena Vani and Bhavya Lahari, An efficient design and implementation of Vedic multiplier in quantum-dot cellular automata, *Telecommunication Systems (TELS)*, 2020.
- [6] X Yang, Y Yu,Estimating Soil Salinity Under Various Moisture Conditions: An Experimental Study, *IEEE Transactions on Geosciences and Remote*, 2017.
- [7] pR Sabia, A Camps, M Talone,Determination of the sea surface salinity error budget in the soil moisture and ocean salinity mission, 2009.
- [8] SA Shahid, M Zaman, L Heng, Soil salinity: historical perspectives and a world overview of the problem, for salinity assessment, mitigation, 2018 – Springer.
- [9] RMA Machado, RP Serralheiro,Soil salinity: effect on vegetable crop growth. Management practices to prevent and mitigate soil salinization, *Horticulturae*, 2017 - mdpi.com
- [10] SW Childs, RJ Hanks Model of soil salinity effects on crop growth, *Soil Science Society of America Journal*, 1975 - Wiley Online Library.
- [11] NKR Beechu and et al., System level fault-tolerance core mapping and FPGA-based verification of NoC, *Microelectronics Journal*, Vol 70, PP. 16-26, 2017.
- [12] NKR Beechu and et al., Hardware implementation of fault tolerance NoC core mapping, *Telecommunication Systems*, Vol. 68, 2018.
- [13] B. Naresh Kumar Reddy, Design and implementation of high performance and area efficient square architecture using Vedic Mathematics, *Analog Integr Circ Sig Process*, 2019.
- [14] B. Naresh Kumar Reddy, M.H.Vasantha, Y.B.Nithin Kumar and Dheeraj Sharma, Communication Energy Constrained Spare Core on NoC, 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-4, 2015.
- [15] Naresh Kumar Reddy Beechu, Vasantha Moodabettu Harishchandra, Nithin Kumar Yernad Balachandra, An energy-efficient fault-aware core mapping in mesh-based network on chip systems, *Journal of Network and Computer Applications*, Vol 105, pp. 79-87, 2017.
- [16] Naresh Kumar Reddy Beechu, Vasantha Moodabettu Harishchandra, and Nithin Kumar Yernad Balachandra.,High-performance and energy-efficient fault-tolerance core mapping in NoC, *Sustainable Computing: Informatics and Systems*, Vol 16, pp. 1-10, 2017.
- [17] Beechu, N., Moodabettu Harishchandra, V. & Yernad Balachandra, N. Energy-Aware and Reliability-Aware Mapping for NoC-Based Architectures, *Wireless Personal Communications* 100 (2), 213-225, 2018.
- [18] B. N. K. Reddy, M. H. Vasantha, Y. B. Nithin Kumar and D. Sharma, A fine grained position for modular core on NoC, 2015 International Conference on Computer, Communication and Control (IC4), Indore, pp. 1-4, 2015.
- [19] B. Naresh Kumar Reddy, M.H.Vasantha and Y.B.Nithin Kumar, A Gracefully Degrading and Energy-Efficient Fault Tolerant NoC Using Spare core, 2016 IEEE Computer Society Annual Symposium on VLSI, pp. 146-151, 2016.
- [20] Sai Kumar, T.V.K.Hanumatha Rao and B. Naresh Kumar Reddy, Exact Formulas for Fault Aware Core Mapping on NoC Reliability, 17th International IEEE India Conference INDICON, 2020.
- [21] B. Naresh Kumar Reddy, Dharavath Kishan & B. Veena Vani, Performance constrained multi-application network on chip core mapping, *International Journal of Speech Technology* , Vol 22(4), pp. 927-936, 2019.
- [22] Reddy B.N.K., B.Sireesha ,An Energy-Efficient Core Mapping Algorithm on Network on Chip (NoC) VLSI Design and Test. VDAT 2018. *Communications in Computer and Information Science*, vol 892, pp. 631-640, Springer, Singapore, 2019.
- [23] VS Boddu, BNK Reddy, MK Kumar, Low-power and area efficient N-bit parallel processors on a chip, 2016 IEEE annual India conference (INDICON), pp. 1-4, 2016.
- [24] B Naresh Kumar Reddy, G Sai Vishal Reddy, B Veena Vani, Design and Implementation of an Efficient LFSR using 2-PASCL and Reversible Logic Gates, *IEEE Bombay Section Signature Conference (IBSSC)*, pp. 247-250, 2020.
- [25] Sudheer H, G Sai Vishal Reddy and B Naresh Kumar Reddy, Design and Analysis of High Reliable Fault Tolerance Subsystem for Micro Computer Systems, 11th IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE 2021), pp. 127-130, 2021

Transmission of Biometric Feature Using Facial Features Securely For Long Distance Biometric Recognition System

M. Suganthy

Associate Professor

Department of Electronics and Communication
Engineering

VelTech MultiTech Dr Rangarajan

Dr. Sakunthala Engineering College
Chennai, India

suganthym46@gmail.com

M. Kavitha

Professor

Department of Computer science Engineering
Veltech Dr Rangarajan Dr Sakunthala R&D
Institute of Science and Technology
Chennai, India

mkavi277@gmail.com

S. Manjula

Associate Professor

Department of Electronics and Communication
Engineering

Rajalakshmi Institute of Technology
Chennai, India

manjulasankar@gmail.com

P. Anandhan

Professor

Department of Electronics and Communication
Engineering

Saveetha school of engineering, Saveetha institute of Medical
and Technical Sciences ,Saveetha University
Chennai, India

anandanvp2000@gmail.com

Abstract- Assuring biometric data security and integrity is crucial. The widespread use of biometric authentication technologies creates a demand for data security and privacy that is both effective and dependable. To solve these difficulties, watermarking approach is preferred for biometric data protection. Multi biometric systems have enhanced the accuracy and reliability of biometric systems. Though a biometric system can be hacked in a variety of ways, unauthorised access to biometric template information poses a severe security and privacy risk. Here description of biometric watermarking strategies for safe user verification on a remote, biometric system that uses both facial and multi iris information, and statistically assess their effects on verification accuracy is presented. The Haar wavelet, LBP, and LTP iris characteristics are combined using average fusion and concatenation fusion. A single secure sketch is created after messing with the images. This biometric feature is hidden within the cover image, facial image. The results of the experiments reveal that the preset's security, performance, and accuracy are all satisfactory.

Keyword - Facial Image, Multi biometric, LBP, Harr Wavelet

I. INTRODUCTION

In communication networks, data transmission security is a critical issue. A communication system is trustworthy if it offers a high level of security. Also the effectiveness of a biometric system, as well as its security against intruders, unauthorised alteration, and misuse, are key factors in its reliability and user acceptance. The system is effective

when it verifies that the templates made of biometric features originated from a valid and authentic user at the time of enrollment. In this instance, the transmission medium should offer security, integrity, authenticity, and secrecy of the transmitted data. Nowadays, online multimedia is increasingly popular; for every second a large amount of data is sent through an unsecure route, which could be dangerous. As a result, it's critical to keep data safe from hackers.

Techniques like as cryptography and steganography are used to secure data. Cryptography encrypts data to ensure secure transmission [8]. Prior to transmission, encryption is employed and the decryption is performed over the encrypted data . The science of embedding hidden messages in other digital information is Steganography ; it sends data by hiding it in a cover item, such as a picture or audio.

Watermarking is a sort of steganography in which encapsulation of one information into another, both are somehow connected. Physical items are protected from counterfeiting and illegal copying using watermarking technology.In our previous works [1-7], [16] various feature extractions for iris biometric system is presented In the construction of a steganography system, various issues are addressed, such as invisibility and security [9]. In [10] proposes a steganographic system that exploits human visual sensitivity to conceal hidden bits. To do so, the secret data is first translated into a set of symbols, which are then inserted in a multi-base otation system. The bases used in this circumstance are determined by the level of local variation in the pixel

magnitudes in the host image.. [11] presented a change to the least significant bit matching (LSBM) steganography. Instead of being random, as in LSBM, this change offers the required option of a binary function of two cover pixels. A hybrid data encoding and hiding procedure was developed in [12] to raise the level of security. This was a lengthy procedure. The problem of image colour shifts after the embedding process was solved using this method. The steganography technique [13] is the way to protect the message, to ensure that it is resistant to image steg analysis based on statistical analysis. [14] suggested a new picture steganography that uses the integer wavelet transform [IWT] to incorporate numerous confidential images and keys in a colour cover image. There are two primary elements in the suggested approach[15]. The first is advanced cryptography, while the second is Steganography is one of them. Here, cryptography is improved since it is distinct from previous cryptography in that it has its own key. Secure communication and patient privacy protection necessitate techniques such as cryptography and watermarking[17]. Video must be compressed and sent in most applications if remote identification system is required, which is the focus of this research.

II. PROPOSED SYSTEM

A. System Design

In this proposed system watermarking is used to secure multi biometric feature template. The system design is depicted in figure 1. Encryption, watermarking, and steganography are viable approaches to strengthen the security of biometric data in order to promote widespread use of biometric systems. Digital watermarking is used to protect iris features.

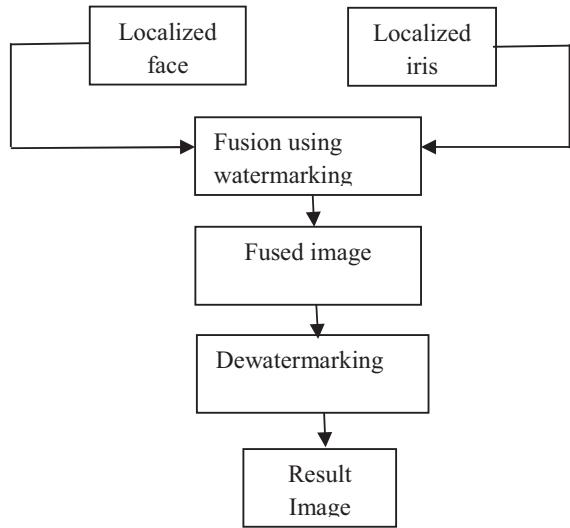


Fig .1 System design

B. Process in Face Detection

The following sections describe the process involved in face detection. Figure 2 shows how the face is detected from the frame. It comprises of the following modules.

- i. Encoder
- ii. Input Quantization
- iii. Extraction of wavelet Coefficient
- iv. DWT
- v. Classification

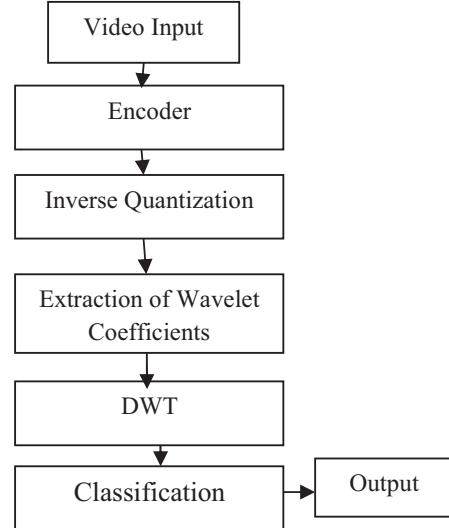


Figure .2 Face detection process

a) Wavelet sub band extraction

Following inverse quantization, the wavelet sub bands of Approximation Low Low (ALL) and Approximation Low High (ALH) of luminance (Y) and ALL of chrominance of the reference video frame are retrieved in the first stage of the proposed system. For face identification, the ALL and ALH of luminance are retrieved. For face localisation, the entire chrominance spectrum is taken. Because the sub bands required for face recognition (ALL, ALH) are level-2 decomposition sub bands, the compression system's degree of decomposition is set to 2. For correctly recovering the ALL and ALH sub bands from video frames, the wavelet decomposition coding used in the compression strategy is important. Figures 3.a and 3.b depict the wavelet decomposition procedure.

This compression scheme employs the lifting wavelet transform, which consists of three phases. The first step is to split the given sequence as even and odd, and then use the even part to predict the odd part in the second step. The difference between the expected odd sequences and the odd sequence is the high-frequency coefficient, abbreviated as H. The low-frequency coefficient, abbreviated L, is generated by combining the even and odd sequences.. The lifting wavelets decomposition is done first for the rows and then for the columns in a 2D image.

ALL	HLL	ALH	HLH
VLL	DLL	VLH	DLH
HL		HH	

a. First level extraction

AALL	HALL	HLL	AALH	HALH	HLH
VALL	DALL		VALH	DALH	
VLL	DLL		VLH		DLH
HL		HH			

b. Second level extraction

Figure. 3 Wavelet subband extraction process

b) Location of face

Determination of face features using color is described in second stage. In video, colours are frequently divided into luminance and chrominance components. To begin, all of the chrominance part sub bands that offer skin colour data are retrieved. A threshold value for the Colour component is used to filter out the area other than skin. Then morphological operations is used to fill the holes which avoids incorrect detection of face region. To develop a threshold value for the chrominance part of skin, the wavelet coefficients values in the chrominance of the ALL wavelet sub band were investigated. The resulting threshold can then be used to differentiate the facial region from the backdrop.

c) Face segmentation

Using the knowledge-based strategy, the face part from the background is separated. This method is based on some simple limitations on the position of the facial region in video frames. Also observed that the aspect ratio of the bounded rectangles of 1:2 was the best acceptable for our video frames examined while using rectangles with specific aspect ratios.

d) Eyes and mouth localization

In this stage location of eyes and mouth is performed to determine the face face features effectively. Depending on the location of eye and mouth , the face is normalized to a standard one.

C. Iris Multi Feature Extraction

With today's world's growing insecurity, where no one knows when their identity card, password, or signature will be stolen, there is a higher demand for identification systems that recognise persons based on difficult-to-steal qualities. With its individual characteristics, the Iris provides a reliable technique of identifying people.Iris

characteristics are extracted using two methods: frequency-based and spatial-based. In the frequency-based technique, the Haar wavelet is utilised to extract features, which are then classified using PCA. Local Binary Pattern(LBP) and Local Ternary Pattern (LTP) are used to extract features in the spatial domain, followed by the estimation of gray Level Co occurrence matrix(GLCM) parameters.

a) LBP Feature Extraction

LBP features are obtained by doing following steps.

1. Divide the window being inspected into cells (e.g. 16x16 pixels for each cell)
2. Compare each pixel in a cell with each of its eight neighbours. Follow the pixels in a clockwise direction around a circle.
3. Write "1" where the value of the centre pixel is greater than the neighbor's. Otherwise, put "0" in the box. This yields a binary number of eight digits.
4. Calculate the frequency of each number occurring in the cell as a histogram..
5. Normalize the histogram as much as possible.
6. All cells' normalised histograms should be concatenated. This returns the window's feature vector.

b) LTP Feature Extraction

The LTP has been suggested as a way to reduce noise dependence, especially near the uniform region. Instead of the two-value encoding [0, 1] used in LBP, LTP uses a defined threshold to generate three encoding values [1, 0, 1]. In this situation, the criterion used is 5.

c) GLCM

Using a statistical method, such as a co-occurrence matrix, can help offer useful information about the relative positions of adjacent pixels in an image. The co-occurrence matrix P can be defined as in Equation 1 given an image I of size NXN.

$$P_{\Delta_x, \Delta_y}(i, j) \sum_{p=1}^n \sum_{q=1}^m = \begin{cases} 1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta_x, q + \Delta_y) = j \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where, $I(p,q)$ is intensity of the centre pixel and matrix $P(i,j)$ is GLCM . The four parameters (features) that are calculated here are,

1. Contrast = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 \cdot g(i, j)$ (3)
2. Correlation : Gives the measure of correlation of a pixel with its neighbour which is given in Equation 4 .

Correlation

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu)(j - \mu) \cdot g(i, j) / \sigma \quad (4)$$

where, σ is a variance and μ is a standard deviation.

3. Homogeneity: The GLCM homogeneity function returns a value that indicates how close the GLCM element distribution is to the GLCM diagonal. Homogeneity is defined as in Equation 5 for a diagonal GLCM.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (5)$$

The suggested system performs the operation by extracting multifeatures (wavelet and spatial) from the iris. Feature level fusion is used to combine various features from the same iris. At the feature level, the current research on wavelet and spatial fusion yielded greater recognition than other level fusions. The goal of this study is to look into combining scores from two iris systems (left and right). The score is then combined with a key to create a secure biometric template.

III. RESULTS AND DISCUSSION

The results of video face are described here. Finally the segmented iris feature is hidden in chrominance of face by using watermarking to secure iris template. The video input is provided to the face recognizing algorithm. It converts the video sequence into 46 frames per second, with a pixel of 640x272. Finally 46 total video frames available for the process. The following are the results that obtained when applying the face recognition algorithm. The Figure 4 represents the Luminance of the face being recognized. It provides illumination related information of the image. The Figure 5 represents the chrominance of the face being recognized. The color related informations can be obtained from this image.



Figure .4 Luminance of input image



Figure. 5 Chrominance of input

Wavelet coefficient values of chrominance shown in Figure 6. This varies according to the type of background as noncomplex and complex background. The Figure 7 represents the luminance of ALL band and Figure 8 represents the chrominance of ALL band.

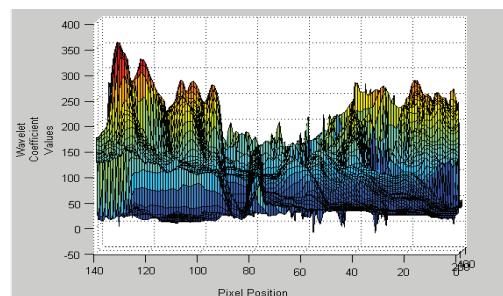


Figure 6 Wavelet coefficients of chrominance



Figure .7 Luminance of ALL band



Figure .8 Chrominance of ALL band

The Figure 9 represents the ALH of luminance of the face being recognized. This image is used as the base image to hide the watermarked image to provide security.

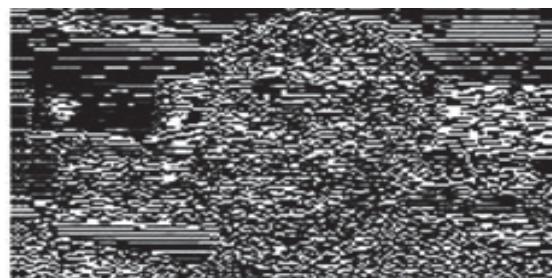


Figure. 10 Watermarked image

The output for fusion techniques, as well as the output for biometric authentication, are displayed. Texture and wavelet feature extraction techniques such as Local Binary Pattern (LBP), Local Ternary Pattern (LTP), Haar wavelet, and the output for biometric authentication are shown. Figure 11 shows the left input image of a human. Figure 12 shows the conversion of a colour image into a grayscale image for image enhancement. For further processing, features are taken from the grey scale image. Figure 13 shows the texture features extracted by LBP, as well as the LTP output images in Figure 14 (a) and (b)

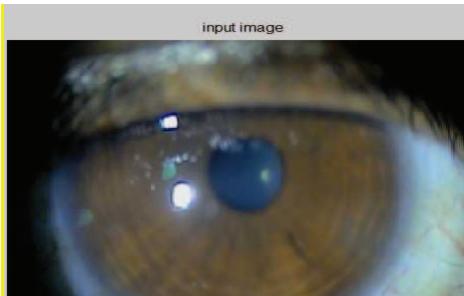


Figure . 11 Input color image

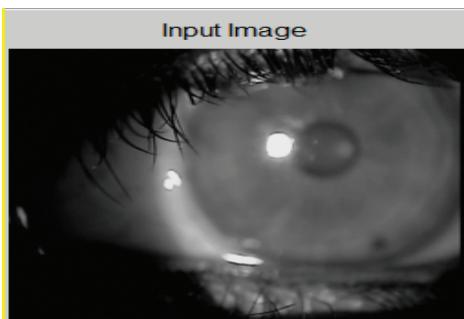


Figure . 12 Gray scale image

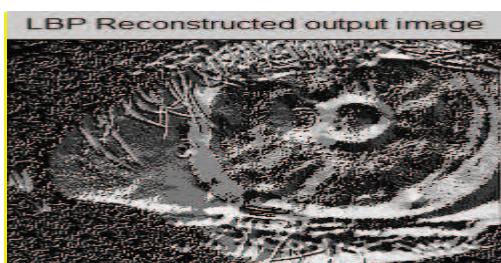
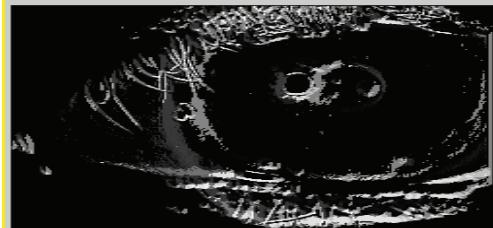


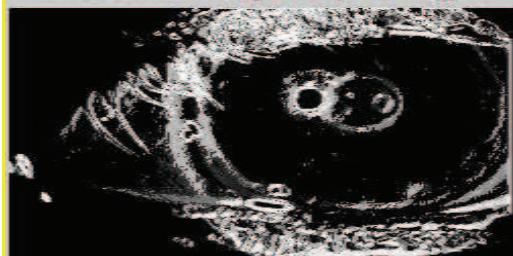
Figure 13 LBP output

LTP upper image output image



a. Upper LTP

LTP lower image output image



b. Lower LTP

Figure 14 Results obtained for LTP

The GLCM parameters are calculated for LBP and LTP. Four parameters are calculated,

Contrast	5.2176
Homogeneity	0.8414
Energy	0.0547
Correlation	0.5913

For the same input image the wavelet features are extracted using Haar wavelet technique and 36 parameters are obtained, shown in Table I. In Table II , the accuracy for various features considered are given.

Table. I Wavelet Features

-0.4021	-0.1722	-0.3021	-0.1410	-0.1763	-0.1899
-0.3052	-0.1415	-0.1899	-0.1379	-0.1722	-0.1763
-0.3021	-0.1379	-0.1763	-0.3021	-0.1415	-0.1722
-0.189	-0.3052	-0.1722	-0.1899	-0.1379	-0.1415
-0.1763	-0.1379	-0.1379	-0.1415	-0.1379	-0.1379
-0.1415	-0.1763	-0.1379	-0.1722	-0.1415	-0.172

The spatial and spectral features are combined using average fusion and concatenation fusion. A single secure sketch is created after messing with the images. This is then buried inside the facial image and kept in a database.

Table II Accuracy

Extracted Feature	Accuracy
LBP	89.4
LTP	90.23
Harr wavelet	92.54
LBP +LTP	95.62
LBP +LTP+ Harr wavelet	98.56

IV. CONCLUSION

The performance of this proposed work is assessed based on the accuracy of identification obtained. As a classifier, the SVM is utilised. Matlab is used to carry out the implementation. The outcomes are expressed in terms of classification accuracy, which is calculated as the percentage of correctly categorised images throughout the testing phase. As a result, the solution is a flexible platform for safeguarding biometric templates. Because of its dependability and precision, iris recognition is becoming more popular. Iris recognition algorithms will be used in a variety of applications if they are tuned for low-cost specialised hardware. Future research in this area could include the use of different biometric along with watermarks and cryptography.

V. REFERENCES

1. Suganthy, M & Ramamoorthy, P 2012, 'Principal component analysis based feature extraction, morphological edge detection and localization for fast iris recognition', Journal of ComputerScience, vol. 8, no. 9, pp. 1428-1433.
2. Suganthy, M & Ramamoorthy, P 2013, 'Novel Feature Extraction for Iris Recognition', European Journal of Scientific Research, vol. 95, no.3, pp.423-429.
3. Suganthy, M & Ramamoorthy, P & Krishnamoorthy, R 2012, 'Effective Iris Recognition For Security Enhancement', International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, vol. 2, no. 2, pp. 1016-1019.
4. Suganthy, M & Ramamoorthy, P 2012, ' An efficient authentication by iris using log Gabor filter and neural network', IJCSI International Journal of Computer Science Issues, vol. 9, no. 1, pp. 371-376.
5. Suganthy, M & Ramamoorthy, P 2011, 'An efficient authentication using fusion of different modalities,' Ciit International Journal of Digital Image Processing, vol. 3, no. 7, pp.410-414.
6. Suganthy, M & Ramamoorthy, P, Veeralakshmi & Suresh, S 2014, 'Voice recognition system for physically disable using spectrogram techniques', Asian Journal of Pharmaceutical Research, vol. 4, no. 3, pp. 112-121.
7. Suganthy, M, Ramamoorthy, P & Kanagalakshmi, A 2014, 'Wavelet and texture feature fusion of dual iris based security system using T-Norm', International Journal of Applied Engineering Research (IJAER), vol.9, no. 24, pp. 30681-30692.
8. Obaida Mohammad Awad Al-Hazaimeh, (2013) "A New Approach for Complex Encrypting and Decrypting Data" International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.2.
9. Katzenbeisser, S. and Petitcolas, F.A.P. 2000, Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, Inc., Boston, London.
10. Xinpeng Zhang and Shuzhong Wang, (2005), "Steganography Using MultipleBase Notational System and Human Vision Sensitivity", IEEE signal processing letters Vol. 12, No. 1.
11. Jarno Mielikainen, (2006), "LSB Matching Revisited", IEEE signal processing letters, Vol. 13, No. 5.
12. Piyush Marwaha, Paresh Marwaha, (2010), Visual Cryptographic Steganography in images", IEEE, 2nd International conference on computing,Communication and Networking Technologies.
13. G.Karthigai Seivi, Leon Mariadhasan and K. L. Shunmuganathan, (2012), "Steganography Using Edge Adaptive Image" IEEE, International Conference on Computing, Electronics and Electrical Technologies.
14. Hemalatha S, U Dinesh Acharya, Renuka A and Priya R. Kamath, (2012), " A Secure and High Capacity Image Steganography Technique", Signal & ImageProcessing : An International Journal (SIPIJ) Vol.4, No.1.
15. Darshana Patil & P. M. Chawan, 2017 "A Secure Data Communication System Using Enhanced Cryptography and Steganography "International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, no.6.
16. M. Suganthy, S Manjula , "Performance analysis of iris biometric system using GKPCA SVM International Journal of InformationTechnology and Management vol.20, 07-216. Inderscience publishers.
17. Madhusudhan, K.N., Sakthivel, 2021 , "A secure medical image transmission Algorithm based on binary bits and Arnold map", J Ambient Intell Human Comput 12, 5413–5420.

Real Time Continuous Behavior Class Based Secure Routing for Improved QoS in Industrial Networks

¹BABU.N, ²TAMILARASI SURESH

¹Research Scholar/CSE,²Professor/IT

^{1,2}St. Peter's Institute of Higher Education and Research,

Chennai, Tamilnadu, India. Pincode: 600054.

¹babuskpt@gmail.com, ²tamilarasisu@gmail.com

Abstract: The problem of data routing in industrial networks has been well studied and analyzed. Towards the scope, there exist numerous techniques in routing data packets in industrial networks which would use several features of network conditions like traffic, congestion and so on. However, they suffer to achieve higher performance in many QoS parameters. In the intention to improve the QoS performance of industrial networks, an efficient Continuous Behavior Class Based Secure Routing (CBCSR) algorithm is presented in this paper. The method not just consider the behavior of different components of the industrial network but also looks on the key metrics of network and how the nodes behave on different conditions. To perform this, the method identifies various routes to reach any destination where the service available including different interfaces, components, devices and other tiny networks. According to the nodes and devices identified, the method discovers routes available as per the geographic location of wireless devices independent of their mobility. Within the list of routes identified, the method computes the value of Class Based Secure Route Measure (CBSRM) for various classes of packets. Finally, a single route has been identified and selected for data transmission. The proposed approach improves the performance in secure routing and support the development of QoS of industrial IoT sensors networks.

Keywords. Industrial IoT Sensors Networks, Secure Routing, QoS performance, CBCSR, CBSRM, Routing Attacks.

1. Introduction

The recent organizations deploy their own networks to perform several processes in an automated manner. Various tiny networks in heterogeneous nature are integrated to form the industrial networks which comprises of different components from devices, sensors, components and networks towards providing accessibility for the service users in reaching the service point. However, the components and devices are dedicated for the access of organization's users to perform variety of operations.

The industrial network has several characteristics which splits the network resources into several levels. Initially, the data network has been completely separated from business network. The data network has several nodes which are interconnected and possess the data maintained by the network. However, there are many controllers and gateway, routers to reach the data servers. Also, each component and device would be geographically distributed. Similarly, the units of industry would be located within a specific region which can be covered by WAN, CAN, LAN or it may be located faraway. However, to provide access to the network resources and the data, there are number of

interconnecting devices like physical cables and wireless devices in the network.

In most case, the network devices and resources are connected through wireless devices. However, the sensor nodes are limited in their physical characteristics like transmission range and energy. The restriction in transmission range limits the nodes in communicating with the set of nodes which are geographically located within the transmission range of the sensor. So, the sensor node has to approach the other nodes in completing the data transmission. This is where the issue comes in the form of network threats. In general, the devices are involved in cooperative transmission to reach the service point or to deliver the data packet to the destination nodes; however, the performance of routing is greatly depending on the way of route selection.

In general, the route selection is performed according to different features and characteristics. For example, the traffic-based approaches use the traffic also the key in the selection of route where the hop-count based approach would use the hop counts of the route in route selection. Similarly, trust-based approaches measure the trust value of nodes in route selection which is being measured according to their behavior in transmission. However, the presence of malicious nodes in the route would affect the throughput performance of routing as well as degrade the QoS performance. There are number of threats which challenge the smooth functioning of industrial network. This paper presents a novel Continuous Behavior Class Secure Routing (CBCSR) scheme which classifies the class of data and class of behavior in routing around the nodes.

2. Related Works

Numerous techniques are analyzed around the problem of secure routing in industrial networks. Such approaches are classified and discussed in this section.

In [1], an efficient biometric based authenticated geographic opportunistic routing scheme is presented to handle IoT devices. The Biometric based-Authenticated Geographic Opportunistic Routing (BAGOR) algorithm depends on the user biometrics to shield the violation of DoS attacks.

In [2], they proposed a hybrid routing protocol which uses existing digital certificate and other cryptographic techniques, is secure and efficient. In [3], a Multi-Attack Detection using Forensics and Neural Network is presented.

In [4], an ant colony based QoS aware energy balancing secure routing (QEBSR) is presented which computes the trust according to the end to end delay.

In [5], the author propose a intrusion prevention framework which forms clusters and performs multi hop routing according to black chain techniques.

In [6], an AODV based approach with modified sequence number algorithm is used to find the black hole nodes in the network.

In[7] A DSOR: A Traffic-Differentiated Secure opportunistic Routing with Game Theoretic Approach is presented, which consider forwarding capability of nodes in measuring the trust.

In [8], a trust-based hybrid secure routing scheme (S-DSR) is presented, which uses the trust value of nodes collected from neighbor nodes in establishing secure route for communication.

In [9],a Novel Intrusion Detection and Prevention scheme is presented with various case studies.

In [10], a new secure protocol Hybrid Cryptography-Based Scheme (HCBS) is presented, which is developed on a combination of the cryptography technique based on Elliptic Curves to exchange keys that uses symmetric keys for data encryption and MAC operations.

In [11], a trust and energy aware secure routing protocol (TESRP) is presented towards wormhole attack. The method uses sequence number concept has been used for securing TESRP from wormhole attack.

In[12] A trust sensing-based secure routing mechanism (TSSRM) with the lightweight characteristics is proposed.

In[13] A Q learning based efficient secure routing (ESRQ) is presented], which computes the trust according to the behavior of the nodes. A trust based secure energy efficient routing (TSER) is presented.

In [14],which select the route according to hop count, energy. The nodes are authenticated based on key exchange and messages.

In[15],trust sensing-based secure routing mechanism (TSSRM) with the lightweight characteristics is proposed.

In[16],A Q learning based efficient secure routing (ESRQ) is presented which computes the trust according to the behavior of the nodes.

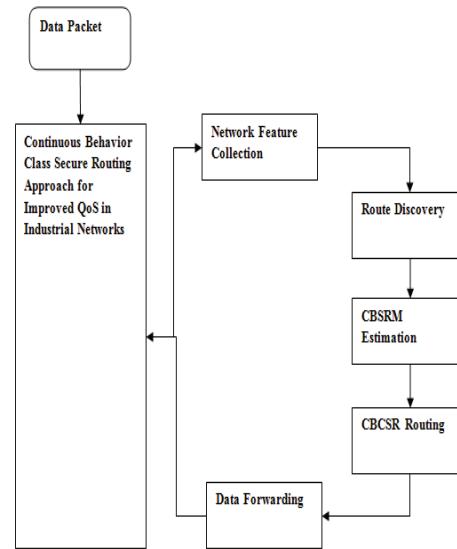
In[17],A Iterative Filtering (IF) algorithm based secure routing is presented to detect the Simple attack and Collusion attack on nodes.

In[18],A trust based secure energy efficient routing (TSER) is presented.

3. Continuous Behavior Class based Secure Routing (CBCSR) Approach

The proposed continuous behavior class based secure routing algorithm collects the network statistics from the topology initially. From the topology constraints, the method finds the set of components, devices, sensors, working units and network routers. Using them, the method discovers the list of routes to reach any service point. Further, for each route discovered with the traces of previous transmission history, the method estimates the class based secure route measure (CBSRM) value. Finally, a route with higher CBSRM has been selected for data transmission. The detailed approach is discussed in this section.

Figure 1. Architecture of Proposed CBCSR Routing Approach



The functional architecture of proposed CBCSR routing algorithm is presented in Figure 1, where the stages involved in the approach is discussed in detail in this section.

3.1. Network Feature Collection:

The routing in industrial network is not just through the routers and sensors, but also depending on other components and devices. To perform this, the method first identifies the set of components, devices, sensors, interfaces, IoT devices present in the network. Using the network topology, the method identifies the location of each component present in the network. Such identified components with their location details with other characteristics are added to the network feature set. Such feature set collected has been used to perform routing in industrial networks.

Algorithm 1:

```

Given: Network Topology NT.
Obtain: Network Feature Set NFS.
Start
Read NT.
Find Components Comp =  $\sum$  Components  $\in$  NT
Find Sensors Sens =  $\sum$  Sensors  $\in$  NT
Find IoTSlots =  $\sum$  IoTs  $\in$  NT
Find Devices Devs =  $\sum$  Devices  $\in$  NT
Find Routers Rs =  $\sum$  Routers  $\in$  NT
For each component, device, sensor, IoT and
Routers
Identify location L.
Identify transmission range Tr.
Identify Traffic Trf.
Identify energy Er.
Feature set Fs = {ID, L, Tr, Trf, Er}
NFS =  $\sum$ (Featureset  $\in$  NFS)  $\cup$  Fs
End
Stop

```

The above discussed algorithm represents how the network feature collection is performed. It finds the location, traffic, transmission range and energy of all the devices or components identified. Generated feature sets are added to the set to be used on secure routing.

3.2. Route Discovery

Route discovery is the process of identifying set of routes between any two source and destination. To perform this, the method uses the network feature set generated. Using the feature set, the method identifies the set of devices or components located within the location of source and destination. Now, within the components identified, the method uses the location detail to fetch the route available. For each component through other components, the method maps the location of various other devices recursively to reach the destination. Identified routes are used in route selection by computing CBSRM value.

Algorithm 2:

```

Given: Network Feature Set NFS, Source S,
Destination D.
Obtain: Route Set Rs.
Start
Read NFS, S, D.
For each component C in NFS.
For each component C1 in NFS
Identify set of routes to reach Destination D.
If
 $\sum_{i=1}^{size(NFS)} \sum_{j=1}^{size(NFS)} Routes \leftrightarrow C, C1$  then
Add to route set Rs =  $\sum$  Routes  $\in$  Cy, C1}
End
End
Stop

```

The above discussed algorithm finds the routes between the source and destination according to the transmission range of various nodes in the network around the source. Identified routes has been used towards secure routing and route selection.

3.3. CBSRM Estimation

The class based secure route measure represent the fitness of any route considered. For example, if a route R is considered, then the value of CBSRM represents the fitness of the route towards successful transmission of data. It has been measured based on the number of mobile sensors, number of high configuration devices, number of IoT devices, number of routers present in the network. Also, it depends on the transmission made by each device or component and their success rate. Similarly, the CBSRM value is depending on the traffic at each hop considered. The value of CBSRM is measured by computing Route Stiff Measure (RSM) which represent the stability of the route, Transmission stability measure (TSM) which represent the stability of route in transmission, Behavior Support Measure (BSM) which depends on behavior of previous transmission. Using all these measures, the method computes the value of CBSRM.

Algorithm 3:

```

Given: Transmission Trace Tt, Route R.
Obtain: CBSRM
Start
Read transmission trace Tt, Route R.
Compute Route Stiff Measure RSM.
 $RSM = \frac{1}{\sum_{Devices \in R} size(R)} \times \left( \frac{\sum_{IoTDevices \in R}}{size(R)} \times 0.3 \right) \times$ 
 $\left( \frac{\sum_{Routers \in R}}{size(R)} \times 0.8 \right) \times \frac{\sum_{Components \in R}}{size(R)}$ 
Compute Behavior Support Measure
BSM.
 $BSM = \frac{\sum_{i=1}^{size(R)} \sum_{j=1}^{size(Tr)} Tr(i) == R(i) \& State == Transmitted / size(Tr(i), R == R)}{size(R)}$ 
Compute Transmission Stability Measure TSM.
 $TSM = \frac{\sum_{i=1}^{size(R)} R(i).Traffic < Th}{size(R)}$ 
Compute CBSRM =  $\frac{RSM}{BSM} \times TSM$ 
Stop

```

The above discussed algorithm computes the class based secure route measure according to different stability value computed on route, behavior and transmission characteristics. Using the values, the method computes the value of CBSRM to perform route selection.

3.4. CBCSR Routing:

The proposed approach performs continuous behavior class secure routing based on the network feature set generated from the network topology. Using the feature set, the method performs route discovery. For each route identified the method computes CBSM value according to the stability of route in transmission, behavior and characteristics. Based on the value of CBSRM, the method selects a most secure route to perform data transmission. Selected route has been used to perform data forwarding.

4.Results

The methods of secure routing in industrial network have been analyzed for their performance under various circumstances. The performances of the routing methods are evaluated at the presence of different number of nodes in the environment. Their performance is measured under different parameters. The details of simulation are presented as below:

Table 1. Details of Simulation

Parameter	Value
Simulator	NS2
Number of Nodes	500
Simulation Time	10 Minutes
Average Energy	100 Joules

The environmental details being used for the performance analysis of various approaches of secure routing in industrial network is presented in this section. The performances of the methods are analyzed according to the parameters such as Secure Routing Performance, Throughput Performance, Packet Delivery Ratio, and Packet Drop Ratio.

Where the routing methods are the following:

- HCBS - Hybrid Cryptography-Based Scheme
- TSSRM - Trust sensing-based secure routing mechanism
- ESRQ - Q learning based efficient secure routing
- BAGOR - Biometric based Authenticated Geographic Opportunistic Routing
- CBCSR - Continuous Behavior Class Secure Routing

- **Secure Routing Performance:**

The secure routing performance of any approach is measured based on the number of threats generated and number of threats identified by different algorithms. It has been measured as follows:

$$\text{Secure Routing Performance} = \frac{\text{Number of Threats Detected Successfully}}{\text{Number of Threats Generated}} \times 100 \quad (1)$$

Table 2 : Analysis on Security Performance

	Security Performance in %		
	100 Nodes	300 Nodes	500 Nodes
HCBS [12]	71	74	78
TSSRM [15]	73	77	79
ESRQ [16]	75	78	80
BAGOR [1]	77	79	82
CBCSR	79	82	85

The performance in security produced by different approaches is measured at different number of nodes in the network. The CBCSR approach improves the security performance up to 85% which is 13% higher than HCBS, 6%

higher than TSSRM, 5% higher than ESRQ and 3% higher than BAGOR algorithm.



Figure 2. Analysis on security performance.

The performance of the methods on security performance is measured and analyzed in Figure 2, which is analyzed at the presence of different number of nodes in the network.

- **Throughput Performance:**

The throughput has been measured as follows:

$$\text{Throughput performance} = \frac{\text{Total bytes delivered}}{\text{Total bytes transferred}} \times 100 \quad (2)$$

Table 3: Analysis on Throughput

	100 Nodes	300 Nodes	500 Nodes
HCBS [12]	73	75	78
TSSRM [15]	73	77	79
ESRQ [16]	75	78	80
BAGOR [1]	77	79	82
CBCSR	79	82	85

The throughput performance achieved by various approaches in different environment conditions are measured and presented in Table 4. The CBCSR approach improves the throughput performance up to 89% which is 11% higher than HCBS, 10% higher than TSSRM, 9% higher than ESRQ and 3% higher than BAGOR algorithms.

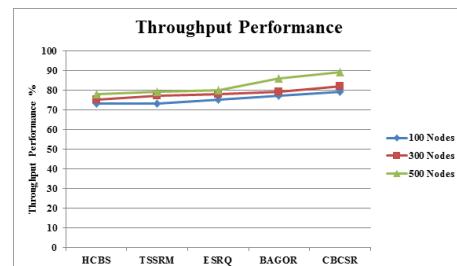


Figure 3. Analysis on throughput performance.

- **Packet Delivery Ratio:**

It has been measured as follows:

$$\text{Packet Delivery Ratio} = \frac{\text{Number of Packets Delivered}}{\text{Total Packets Sent}} \times 100 \quad (3)$$

Table 4: Analysis on Packet Delivery Ratio

	Packet Delivery Ratio in %		
	100 Nodes	300 Nodes	500 Nodes
HCBS [12]	73	75	78
TSSRM [15]	73	77	79
ESRQ [16]	75	78	80
BAGOR [1]	77	79	86
CBCSR	79	82	89

The packet delivery ratio achieved by various approaches in different environment conditions are measured and presented in Table 5. The CBCSR approach improves the packet delivery ratio performance up to 89% which is 11% higher than HCBS, 10% higher than TSSRM, 9% higher than ESRQ and 3% higher than BAGOR algorithms.

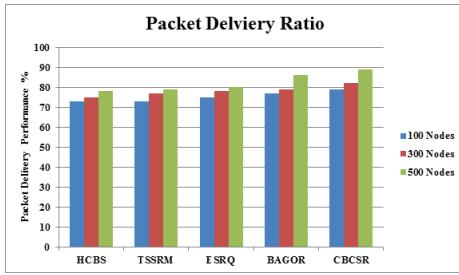


Figure 4: Analysis on Packet Delivery Ratio

Packet Drop Ratio:

The packet drop ratio is the measure which represents the performance of the protocol in delivering the packets. It has been measured according to the number of packets being dropped for a given number of packets sent. It is measured as follows:

$$\text{Packet Drop Ratio} = \frac{\text{Number of Packets Dropped}}{\text{Total Packets Sent}} \times 100 \quad (4)$$

Table 5 : Analysis on Packet Drop Ratio

	Packet Drop Ratio in %		
	100 Nodes	300 Nodes	500 Nodes
HCBS [12]	29	26	22
TSSRM [15]	27	23	21
ESRQ [16]	25	22	20
BAGOR [1]	23	21	18
CBCSR	21	18	15

The ratio of packet drop produced by different methods at the presence of various numbers of nodes in the network has been measured and presented in Table 6. The proposed CBCSR approach has produced less packet drop ratio up to 15% which is 7% less than HCBS, 6% less than TSSRM, 5% less than ESRQ, 3% less than BAGOR approaches.

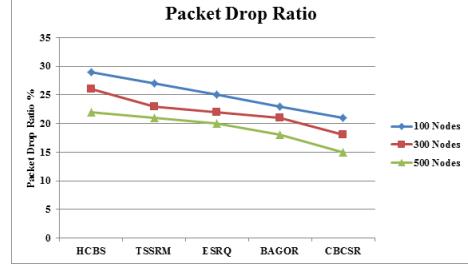


Figure 5: Analysis on Packet Drop Ratio

5. Conclusion

This paper presented a novel continuous behavior class secure routing CBCSR algorithm towards maximization of QoS performance of industrial networks. To perform this, the method generates the feature set from the topology of network and discovers the routes available accordingly. Further, the method computes the value of CBSRM (Class based secure route measure) and based on that a route with higher security is selected. Selected route has been used to perform data forwarding. The method improves the performance of secure routing in industrial network to maximize the QoS performance and Throughput performance also improved. Comparatively packet delivery ratio is very high and packet drop ratio is much low than other methods.

In Future, the QoS performance of the network can be improved by adapting transmission behaviors to perform secure routing in industrial networks.

References

- [1] S.Menaga, An efficient biometric based authenticated geographic opportunistic routing for IoT applications using secure wireless sensor network, SCIENCE DIRECT (MP), 2021.
- [2] K. Biswas and M. Dasgupta, "A Secure Hybrid Routing Protocol for Mobile Ad-Hoc Networks (MANETs)," IEEE (ICCCNT), 2020, pp. 1-7.
- [3] G. Vaseer, "Multi-Attack Detection using Forensics and Neural Network based Prevention for Secure MANETs," IEEE (ICCCNT), 2020, pp. 1-6.
- [4] M. Rathee, S. Kumar, A. H. Gandomi, K. Dilip, B. Balusamy and R. Patan, "Ant Colony Optimization Based Quality of Service Aware Energy Balancing Secure Routing Algorithm for Wireless Sensor Networks," IEEE (TEM), Volume. 68, Number. 1, pp. 170-182, 2021.
- [5] K. Haseeb, N. Islam, A. Almogren and I. Ud Din, "Intrusion Prevention Framework for Secure Routing in WSN-Based Mobile Internet of Things," IEEE, Volume. 7, pp. 185496-185505, 2019.
- [6] S. Shrestha, R. Baidya, B. Giri and A. Thapa, "Securing Blackhole Attacks in MANETs using Modified Sequence Number in AODV Routing Protocol," IEEE (IEECON), 2020, pp. 1-4.
- [7] X. Zhong, R. Lu, L. Li, X. Wang and Y. Zheng, "DSOR: A Traffic-Differentiated Secure opportunistic Routing with Game Theoretic Approach in MANETs," IEEE (ISCC), 2019, pp. 1-6.
- [8] V. S. Bhargavi, M. Seetha and S. Viswanadharaju, "A trust based secure routing scheme for MANETS," IEEE (Confluence), 2016, pp. 565-570.
- [9] G. Vaseer, G. Ghai and D. Ghai, "Novel Intrusion Detection and Prevention for Mobile Ad Hoc

- Networks: A Single- and Multiattack Case Study," IEEE (CEM), volume 8, number 3, pp. 35-39, 2019.
- [10] C. Deepa and B. Latha, "An Energy Efficient Secure Routing (EESR) using Elliptic Curve Cryptography for Wireless Sensor Networks," IEEE (ICICCT), 2018, pp. 1603-1608.
 - [11] R. Shukla, R. Jain and P. D. Vyavahare, "Combating against wormhole attack in trust and energy aware secure routing protocol (TESRP) in wireless sensor network," IEEE (RISE), 2017, pp. 555-561.
 - [12] D. Qin, S. Yang, S. Jia, Y. Zhang, J. Ma and Q. Ding, "Research on Trust Sensing Based Secure Routing Mechanism for Wireless Sensor Network," in IEEE Access, volume 5, pp. 9599-9609, 2017.
 - [13] G. Liu, X. Wang, X. Li, J. Hao and Z. Feng, "ESRQ: An Efficient Secure Routing Method in Wireless Sensor Networks Based on Q-Learning," IEEE (TrustCom/BigDataSE), 2018, pp. 149-155.
 - [14] G. M. Navami Patil, Trust Model for Secure Routing and Localizing Malicious Attackers in WSN, Springer Link (CNS), Volume 12, 2017, PP 1-9.
 - [15] R. Ashtikar, D. Javale and S. Wakchaure, "Energy Efficient & Secured Data Routing Through Aggregation Node in WSN," IEEE (ICCUBE), 2017, pp. 1-6.
 - [16] G. Liu, X. Wang, X. Li, J. Hao and Z. Feng, "ESRQ: An Efficient Secure Routing Method in Wireless Sensor Networks Based on Q-Learning," IEEE (TrustCom/BigDataSE), 2018, pp. 149-155.
 - [17] G. M. Navami Patil, Trust Model for Secure Routing and Localizing Malicious Attackers in WSN, Springer Link (CNS), Volume 12, 2017, PP 1-9.
 - [18] Hilmie Lazrag, A Game Theoretic Approach for Optimal and Secure Routing in WSN, Springer Link (AECIA), Volume 565, 2016, pp 218-228.

Opinion Mining Models for Learner Feedback on Massive Open Online Courses

Christy Jackson J

*School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
christyjackson.j@vit.ac.in*

Suriyakrishnan Sathish

*School of Computer Science & Engineering
Vellore Institute of Technology
Chennai, India
suriyakrishnan.s2019@vitstudent.ac.in*

Aravinda Boovaraghavan

*School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
aravinda.b2019@vitstudent.ac.in*

Narendra G O

*School of Computer Science & Engineering
Vellore Institute of Technology
Chennai, India
narendra.go2019@vitstudent.ac.in*

Maheysh V

*School of Computer Science & Engineering
Vellore Institute of Technology
Chennai, India
maheysh.v2019@vitstudent.ac.in*

Makesh Srinivasan

*School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
makesh.srinivasan2019@vitstudent.ac.in*

Hashwanth S

*School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
hashwanth.s2019@vitstudent.ac.in*

Abstract—Feedback and rating in MOOC's (Massive Open Online Courses) are critical components in the creation of courses as well as a successful learning process. These data can be evaluated to discover any probable patterns and, depending on the input, the course is likely to be improved. Recent research in sentiment analysis has shown that it may be used to identify emotions. Positive and negative feedback classification can be used as an initial parameter to determine the learner's degree of satisfaction. The major goal of our work includes investigating various machine learning methods for sentiment analysis utilizing existing MOOC course feedback as a case study. The emotions shown in the feedback might give information about the learner's degree of pleasure. Logistic Regression, KNN, Naïve Bayes, Decision Tree and Random Forest were the methods adopted in the research.

Keywords—MOOC, Sentiment Analysis, Classification, Natural Language Processing, Bag of Words

I. INTRODUCTION

The advent of technology has profoundly transformed the educational industry. The various methods in which technology could improve learning are an emerging field of interest for researchers. Massive Open Online Courses (MOOCs) [1] ignited focus into the education sector. Widely popular and proven MOOC learning sites include Coursera [2], Udacity,

edX [3], Australia's Open2Study, and UK Futurelearn. MOOC focuses on building classroom environments that are freely available by unrestricted involvement and online programs. The capabilities entail open access, customizability, flexible content verification, accessible infrastructure, and academic objectives. MOOCs give learners help from virtual communities and resources for online contact with the course teachers [4].

Based on opinions taken from reviews of students, earlier research shows that the most significant consideration for students is who is delivering the course [5]. But understanding what can be improved in the course for a given MOOC to be taught by the same team of instructors is more consequential. Recent studies on the use of social media have demonstrated that several psychological intuitions can be identified and analysed by emotion analysis. An early step in evaluating MOOC's emotions includes determining if the response is positive or negative. It should offer an overview of how learners feel about the course. This leads to better decisions to improve learner experience and satisfaction, which is critical in maintaining the MOOC's performance [6]. Additionally, there is no straightforward solution in MOOCs to address sentiment analysis, nor a study of available techniques.

This research provides a detailed comparison of various machine learning methods that are applied to the feedback messages for polarity identification. Therefore, as the overall goal is to identify patterns based on the research, a dataset from an existing MOOC is taken and studied to diagnose habits that enrolled users exhibit based on their emotions

II. RELATED WORK

Analysis of emotions is used to describe the disparate types of behavior. In specific, various methods will classify certain states or patterns, like process mining [7] or conversation analysis. Discussion research takes advantage of the prediction system used in this research to infer positive emotions. Apart from forecasting affective states in MOOCs, related methods are used for forecasting dropouts, determining whether the learner can complete the course (or whether the learner can obtain a certificate), or the rating that the student will achieve. Several independent variables were used in some situations, mainly related to system utilisation (e.g. inactivity times [8]), group participation, video-observation activities [9], and previous task performance. In [10], the authors have predicted average assignment marks and observed the greatest association between the number of prior quizzes attempted and the ranking. Sinha and Cassel [11] have projected grades and graded them into multiple categories: low, medium, moderate, and very moderate.

Mackness et al. [12] questioned how to build a MOOC that would provide meaningful outcomes for the participants. Much of the previous work associated with this topic included collecting data and interviews [13], [14]. In comparison, predictive text interpretation, image extraction, and data mining techniques were used in some previous e-learning work to extract views from user-generated material, such as comments, forums, or blogs [15]–[17]. Ideology is necessary to optimise as it has been shown that learners with a good attitude are more inspired in E-learning settings [18]. Accordingly, earlier research shows that frustration is associated with impaired performance and behaviour. Conversely, dissatisfaction was less related to worse learning [19]. Depending on user-generated textual comments published during the classes, Adamopoulos [20] conducted sentiment analysis to gather the opinions of students on MOOC features such as the course features and institution features. The aim in that study was to decide which variables influence the success of the course, and it is, therefore, necessary to answer the similar but separate issue of what can be changed when the course is presented again. Piryani, Madhavi, and Singh [21] researched Sentiment Mining and Sentiment Analysis studies, and their findings showed that machine learning (supervised) methods prevailed (67.2%) over (unsupervised) lexicon techniques. Moreover, comments (e.g. film comments) were the most common data types, followed by tweets and news stories. All controlled instances and unsupervised objects shall be identified as points of reference. Authors in [22] evaluated three (supervised) ML algorithms namely Naïve Bayes, Maximum Entropy Classification, and Support Vector Machines (SVM). The reason for choosing

these algorithms is that they had different concepts were employed in earlier classification studies. We considered SVM to be the highest accuracy in that analysis but acknowledged the fact that the problem of classifying feelings has higher complexity than few other classification techniques since this message might sometimes be communicated in a sarcastic language, and the preprocessing stage has some subjectivity. Likewise, Boiri and Moens [23] utilised unigrams to be using the same techniques as binary functions (the occurrence or non-occurrence of the feature). Turney [24] has described, in an unsupervised technique, comments using the difference between words "excellent" and "poor," with accuracies ranging from 0.66 for movie reviews to 0.84 of automobile reviews. However, Hu and Liu [25] made a genius intervention in calculating the directionality of the phrases by remembering the words of opinion (adjectives) and scanning for them in positive and negative terms dictionary definitions, taking into consideration the possible effect of negative terms. While there are many methods to execute sentiment analysis but there are only a handful of submissions in MOOCs that provide an assessment of the method. This article will pave the way for a novel innovation in the field and will include a summary of the various methods of machine learning for the classification of MOOC Courses input.

III. DATASET

While the datasets contains reviews of courses from popular online MOOC websites such as Coursera. The attributes are described in Table I.

TABLE I
ATTRIBUTE DESCRIPTION

S.No	Attribute Name	Attribute Data Type	Attribute Description
1	Courseld	Continuous	Unique identifier for a review.
2	Review	Discrete	Assessment of the course given by the learner.
3	Label	Discrete [1-5]	Rating of the course given by the learner in the range of 1-5 (very negative-very positive)

IV. APPROACH

The sentiment analysis carried out in this study focuses on assessing the polarity of the sentiments of the learners in feedback, i.e. whether these are positive or negative feedback. The approach followed in this study involves Natural Language Processing (NLP) and various classification models.

A. Data Preprocessing

1) *Non-Alphanumeric Characters:* Alphanumeric characters are retained whereas non-alphanumeric characters are removed.

2) *Stemming and Lemmatization:* Reducing the word to its root stem brings uniformity to the words used in the text. However, stemming may reduce a word to meaningless form and lemmatization is adhered to, for production of uniform meaningful root words.

3) Stopwords Removal: Removal of unwanted words like ‘a’, ‘the’ etc. from the text, as they do not impact the classification.

4) Lowercasing: Lowering the case of all the words in the text reduces the dimensions of the matrix representation of the bag of words (BoW) model.

B. Bag of Words

Algorithms can be applied only to numbers and hence the preprocessed text and its tokens is converted to the number format using bag of words model, represented by count vectorization, counting the number of occurrences of each word.

C. Classification Models

The bag of words model produces vectors with binary values, which are fed into different classification models, as an experiment to achieve better and meaningful accuracy. Models included in the research are:

1) Logistic Regression: Logistic regression is a binary classifier, which takes binary values like 1 or 0. Sigmoid function (1) is used for classifying the output. Each class is considered individually for multi-class classification, using One versus All principle.

$$y = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (1)$$

where y is the independent variable to be classified, x is the dependent variable, β_0 is the bias term and β_1 is the coefficient term for every single input x .

The sigmoid function takes values from 0 to 1 and if y is more than 0.5, the output class is 1 else 0, concordant with probability rules.

2) K-Nearest Neighbors Algorithm: K-NN is a lazy learner algorithm, which stores the dataset, followed by taking action on the dataset. It assumes similarity between the new data points and available data, using various distance measures like Euclidean distance, Manhattan distance etc.

3) Naïve Bayes Algorithm: It is a probabilistic algorithm based on Bayes’ theorem with an assumption of independence among independent variables.

$$P(C|x) = \frac{P(x|C) P(C)}{P(x)} \quad (2)$$

where $P(C|x)$ is the probability of the output to be class C , given x , $P(x|C)$ is the probability of the output to be x , given class C . $P(C)$ and $P(x)$ are the prior probabilities of the class C and input x .

The class to which the given input will be classified as per the posterior probability of the classes. A given point gets assigned to the class with the highest posterior probability.

4) Decision Tree Algorithm: This is a tree based graph used to obtain every possible solution, based on the constraints provided.

5) Random Forest Algorithm: Random forest algorithm is an extensive model of decision tree where multiple decision trees are used.

V. RESULTS AND DISCUSSION

Classification models were evaluated using metrics like precision, recall and F1-score.

TABLE II
MODEL ACCURACY

S.No.	Algorithm	Accuracy (%)
1	Logistic Regression	75.4
2	Random Forest	74.9
3	K-NN	73.6
4	Decision Tree	69.1
5	Naïve Bayes	27.9

Fig.1 of the distribution of ratings by linear regression shows that 80% of the reviews were very positive. 14% were positive reviews while neutral, negative and very negative account to a minimal percentage. This indicates that the course they took was highly useful and satisfactory for the users.

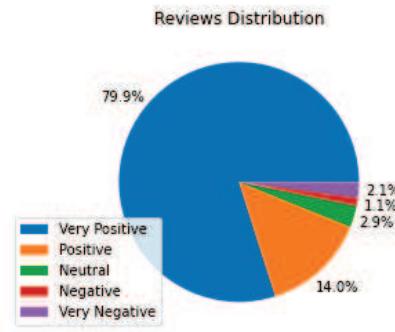


Fig. 1. Distribution of reviews (Logistic Regression).

Fig.2 compares the actual review and reviews predicted by the logistic regression model.

Fig.3 indicates that the computer science courses were inherently popular over the other courses during the most popular courses with its framework courses-Python and R among the top 10. Beyond the top 10 courses, an equilateral trend of technical and non-technical courses were observed.

The word cloud in Fig.4 shows the most popular terms used in Python-based courses, as predicted by logistic regression. User sentiment is largely positive with words such as ‘good’, ‘great’ and ‘learn’ taking part in the reviews. However, certain words, nominal in numbers, like ‘difficult’ and ‘challenging’ indicate negative and neutral sentiments.

Algorithm accuracy can be improved, with improvement in the pre-processing NLP techniques followed by stochastic algorithms like logistic regression, support vector machines, multi-layer perceptrons, with words encoded as in

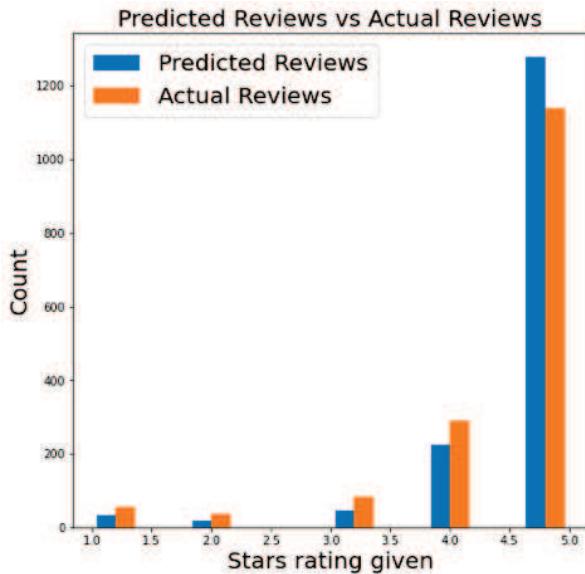


Fig. 2. Predicted vs Actual Reviews (Logistic Regression).

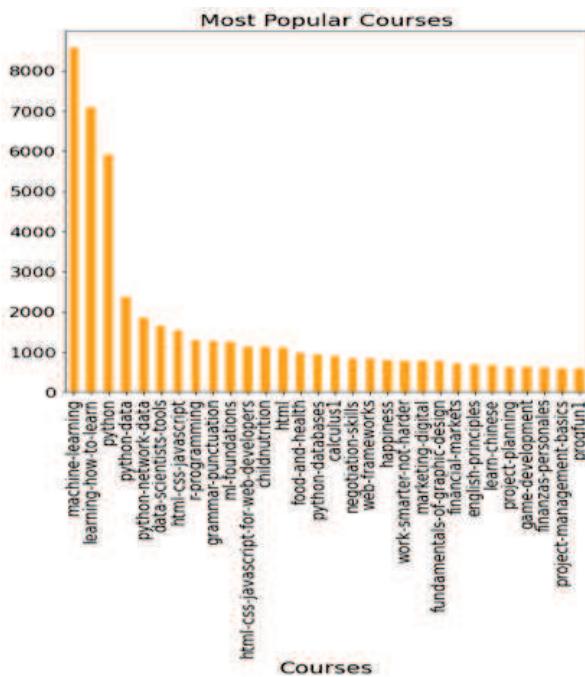


Fig. 3. Popularity of Courses (Logistic Regression).

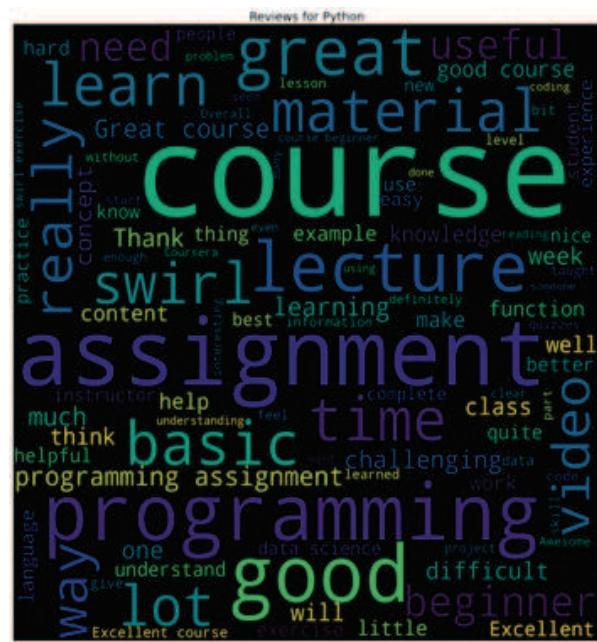


Fig. 4. Word Cloud for Python-based courses (Logistic Regression).

the proposed system, using n-gram or TF-IDF models 3. Classification thresholds for individual categories can also be experimented with. Values above the threshold can be classified as YES (1) and below the threshold can be classified as NO (0).

Since, the ratio of number of samples to the number of words per sample (W) < 1500 , small multi-layer perceptrons with n-gram input can perform better or equivalent to sequence models, with equivalent computation time. If $W > 1500$, fine-tuned pre-trained embedding with sepCNN model provides expected results. Hyperparameter tuning, if applicable, can also be experimented with, changing the number of layers, number of units per layer, dropout rate, learning rate, kernel size, embedding dimensions etc.

$$w = \frac{\text{NumberofSamples}}{\text{NumberofWordsperSample}} \quad (3)$$

VI. CONCLUSION

In this study, various opinion mining models were used to identify the polarity of the course reviews. Patterns with respect to reviews and ratings were obtained, which can help improvement of the courses offered to the learners. Logistic regression achieved the highest comparable accuracy, indicating majorly positive sentiment or 5-star ratings. Our findings have given an insight into the domain of online learning, and effectiveness of MOOC courses among the general population. This is crucial, especially now because MOOC courses have become a primary source of knowledge and in some cases even forms of formal education. Learning about the impacts,

success and effectiveness can help us better structure our programs, and provide better recommendations to learners.

REFERENCES

- [1] J. Kay, P. Reimann, E. Diebold, and B. Kummerfeld, "Moocs: So many learners, so much potential..." *IEEE Intelligent systems*, vol. 28, no. 3, pp. 70–77, 2013.
- [2] C. Severance, "Teaching the world: Daphne koller and coursera," *Computer*, vol. 45, no. 8, pp. 8–9, 2012.
- [3] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. T. Seaton, "Studying learning in the worldwide classroom research into edx's first mooc." *Research & Practice in Assessment*, vol. 8, pp. 13–25, 2013.
- [4] J. Kim, P. J. Guo, D. T. Seaton, P. Mitros, K. Z. Gajos, and R. C. Miller, "Understanding in-video dropouts and interaction peaks in online lecture videos," in *Proceedings of the first ACM conference on Learning@ scale conference*, 2014, pp. 31–40.
- [5] C. Shi, S. Fu, Q. Chen, and H. Qu, "Vismooc: Visualizing video clickstream data from massive open online courses," in *2015 IEEE Pacific visualization symposium (PacificVis)*. IEEE, 2015, pp. 159–166.
- [6] C. G. Brinton, S. Buccapatnam, M. Chiang, and H. V. Poor, "Mining mooc clickstreams: Video-watching behavior vs. in-video quiz performance," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3677–3692, 2016.
- [7] S. Mak, R. Williams, and J. Mackness, "Blogs and forums as communication and learning tools in a mooc," in *Proceedings of the 7th International Conference on Networked Learning 2010*. University of Lancaster, 2010, pp. 275–285.
- [8] T. Phan, S. G. McNeil, and B. R. Robin, "Students' patterns of engagement and course performance in a massive open online course," *Computers & Education*, vol. 95, pp. 36–44, 2016.
- [9] C. Dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 2014, pp. 69–78.
- [10] M. Wen, D. Yang, and C. Rose, "Sentiment analysis in mooc discussion forums: What does it tell us?" in *Educational data mining 2014*. Citeseer, 2014.
- [11] D. Leony, P. J. M. Merino, J. A. R. Valiente, A. Pardo, and C. D. Kloos, "Detection and evaluation of emotions in massive open online courses." *J. UCS*, no. 5, pp. 638–655, 2015.
- [12] S. Halawa, D. Greene, and J. Mitchell, "Dropout prediction in moocs using learner activity features," *Proceedings of the second European MOOC stakeholder summit*, vol. 37, no. 1, pp. 58–65, 2014.
- [13] C. Ye and G. Biswas, "Early prediction of student dropout and performance in moocs using higher granularity temporal information," *Journal of Learning Analytics*, vol. 1, no. 3, pp. 169–172, 2014.
- [14] Z. Ren, H. Rangwala, and A. Johri, "Predicting performance on mooc assessments using multi-regression models," *arXiv preprint arXiv:1605.02269*, 2016.
- [15] T. Sinha and J. Cassell, "Connecting the dots: Predicting student grade sequences from bursty mooc interactions over time," in *Proceedings of the second (2015) ACM conference on learning@ scale*, 2015, pp. 249–252.
- [16] J. Mackness, S. Mak, and R. Williams, "The ideals and reality of participating in a mooc," in *Proceedings of the 7th international conference on networked learning 2010*. University of Lancaster, 2010, pp. 266–275.
- [17] C. O. Rodriguez, "Moocs and the ai-stanford like courses: Two successful and distinct course formats for massive open online courses." *European Journal of Open, Distance and E-Learning*, 2012.
- [18] C. Alario-Hoyos, M. Pérez-Sanagustín, C. Delgado-Kloos, H. A. Parada G., M. Muñoz-Organero, and A. Rodríguez-De-Las-Heras, "Analysing the impact of built-in and external social tools in a mooc on educational technologies," in *European Conference on Technology Enhanced Learning*. Springer, 2013, pp. 5–18.
- [19] D. Song, H. Lin, and Z. Yang, "Opinion mining in e-learning system," in *2007 IFIP international conference on network and parallel computing workshops (NPC 2007)*. IEEE, 2007, pp. 788–792.
- [20] H. H. Binali, C. Wu, and V. Potdar, "A new significant area: Emotion detection in e-learning using opinion mining techniques," in *2009 3rd IEEE international conference on digital ecosystems and technologies*. IEEE, 2009, pp. 259–264.
- [21] A. El-Halees, "Mining opinions in user-generated contents to improve course evaluation," in *International conference on software engineering and computer systems*. Springer, 2011, pp. 107–115.
- [22] J. Moshinskie, "How to keep e-learners from e-scaping," *Journal of Interactive Instruction Development*, vol. 14, no. 1, pp. 8–11, 2001.
- [23] R. S. Baker, S. K. D'Mello, M. M. T. Rodrigo, and A. C. Graesser, "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments," *International Journal of Human-Computer Studies*, vol. 68, no. 4, pp. 223–241, 2010.
- [24] P. Adamopoulos, "What makes a great mooc? an interdisciplinary analysis of student retention in online courses," 2013.
- [25] R. Piryani, D. Madhavi, and V. K. Singh, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015," *Information Processing & Management*, vol. 53, no. 1, pp. 122–150, 2017.

Automatic Stress Recognition System with Deep Learning using Multimodal Psychological Data

1. Mr.Praveenkumar.S,

Research Scholar, Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology,
Chennai, India,
vn8049@srmist.edu.in

2. Dr.T.Karthick,

Assistant Professor, Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology,
Chennai, India,
karthict@srmist.edu.in

Abstract

There are numerous applications for detecting stress states from physiological inputs. Human stress detection may be used to enhance the human experience as well as monitor and prevent stress-related disorders. Previous research has leveraged existing machine learning techniques for automated stress identification. However, there are also some limitations. Deep learning allows you to discover basic patterns in your body's reactions that are often difficult to notice. This study offers a convolutional network-based approach for detecting human stress utilizing multimodal data sets obtained by smart sensors for physiological and motion, which can aid humans in avoiding health issues caused by stress. Sensor information like triaxial accelerates, heartbeat, body temperature, respiration, electromyographic (EMG) data for physiological states stressed and non-stressed, are included in the WESAD data collection. The use of Long short - term memory method and the DNN method to detect stress in people in real-time is investigated in this study. Our research is based on their physical reactions. The study's findings provide an assessment of both models' ability to predict real-time users' mental expressions based on physiological parameter fluctuation.

Keywords-stress, deep neural network, LSTM, DNN, accelerometer, sensors, WESAD

I.INTRODUCTION

Living in society exposes us to a variety of scenarios that necessitate human connection. Social psychology is the study of social processes that have an impact on human behavior. The fields of social psychology and physiological signal analysis are combined in social psychophysiology. In other words, it

looks into the connection between social conduct and the physiological responses that accompany it. Social stress, which originates from a person's resistance to social situations, is one of the topics that social psychology is interested in.

Nowadays, one of the most important areas of research is equipping electronic gadgets with the ability to recognize and classify human stress. Several studies have been conducted utilizing various modalities to detect and classify stress [1]. Some of the modalities associated with stress perception are individuals' physiological, psychological, and bodily responses. Physiological signals are used in a variety of therapeutic applications, including the assessment and classification of mood and stress [2]. Physiological signs such as triaxial acceleration, heartbeat, body temperature, respiration, electromyographic (EMG) data have shown to be fairly accurate in detecting stress [3].

Physiological signals above mentioned are used to quantify human mood [4]. Several studies have used subjective stress assessments (e.g., stress questionnaires) to assess people's stress levels [5]. Stress is directly linked to psychological states, and it has an impact on one's work performance, emotions, mental, and physical health [6]. People who have been diagnosed with severe stress have also reported poor overall health. Furthermore, they have a high incidence of chronic disorders, like high cholesterol, increased blood pressure, overweight, asthma, and diabetes, which put people at risk of dying early [7].

Previous research has also found connections between depression and anxiety and cardiovascular and cerebrovascular illnesses [8].

In recent years, business and academic [9], including human-computer interaction, education, voice recognition, and medical sectors, have placed high importance on stress recognition. It has used expressive modalities such as facial expressions [10], speech [11], body gestures [12], and other explicit features to carry out typical ways for recognizing stress. These methods, however, have limitations in that they cannot evaluate interior feelings and it is difficult to eliminate the influence of subjective elements like phony expression. People can, for example, smile while they are self-deprecating or even melancholy. They may, however, maintain a poker face when they are joyful.

In [13], they have looked towards automated stress detection utilizing wearable gadgets. However, most existing research focuses on developing machine learning algorithms that classify stress based on physiological information collected from wearable sensors. Some approaches classify stress into distinct groups, but they do not allow for the measurement of effective stress or range. Due to a lack of quantitative data, these stress detection approaches may not be able to track the changes in affect scores or provide inter-subject changes over time. Other research has sought to quantify the intensity of emotions like distress [14] as well as anxiety [15].

Many biosignal-based approaches have been investigated for stress perception. However, to track the biosignal, the biosignal measuring device must be connected to the body, so the user may feel rejected. As a result, several studies on stress recognition using thermal imaging have been conducted. However, it is impossible to perceive stress in everyday life because, without a thermal imager, stress cannot be perceived. On the other hand, in the case of stress recognition using wide images, a rather simple function was used in most studies.

II.RELATED WORKS

Early stress detection studies employed biosignals since they capture the most sensitive physiological changes and can identify changes that aren't reflected by the face or behavior. In these investigations, signal-based elements like an

electrocardiogram, respiration, heart rate, and blood pressure were gathered and used to communicate stress. Many of them employ conventional classifiers such as Svm Classifier, Hidden Markov Models, Adaboost, and K - nearest neighbor.

Because when an individual is agitated, their blood flow and facial warmth increase. Many investigations have been carried out to determine the change utilizing thermal imaging. These studies used a variety of approaches for detecting stress, including extracting variables like respiration rate, blink rate, galvanic skin, and blood circulation directly from thermal pictures.

When a person is under stress, the movements of the eyes, mouth, and head are different from the normal state, and studies on stress recognition using a wide range of images are being conducted. In this study, stress was detected using various methods such as eye size, mouth movement, and head movement, which are the functions and characteristics of the hand-extracted from the eye, nose, and mouth regions.

G. Giannakakis et al. [16] created a system to detect and evaluate stress/anxiety emotional states using video-recorded facial signals. Oral activity, ocular events, camera-assisted photo flow measurements, heart rate measurements, and head movement parameters were also assessed. Participants were instructed to sit 30 feet apart before a desktop computer with an integrated camera. The Widely Used Algorithms like Naive Bayes classifier, SVM, K-nearest neighbors, and AdaBoost classifier were all utilized and assessed. The AdaBoost classifier produced the best classification results in the social exposure procedure, with an accuracy of 91.68 percent.

Jacqueline Wijsman, et al. [17] used wearable sensors to assess physiological signals to detect mental stress. They took the subjects' ECG, galvanic skin, respiration, EMG and estimated a total of 19 physiological characteristics from them. After studying correlations and normalizing feature values, a subset of 9 features was selected for further analysis, which was then reduced to 7 features using Principal component analysis. Using these variables and additional classifiers such as the Linear Bayesian Norm Classifiers, Quadratic Bayes Normal Classification algorithm, K-Nearest neighbor Neighbors Classifier, and Fisher's Least Square Linear

Classifier, an accuracy of 80% was achieved among stress and non-stress circumstances of a person.

Saskia Koldijk et al. [18] created a new multimodal dataset called SWELL Knowledge Work (SWELLKW) for stress studies and user modeling. This data was collected while 25 participants performed common tasks such as writing, reading, searching, etc., and changing their work environment through two stressors: lack of time and interruption of email. Physical poses, facial expressions, computer records, skin conductance, and heart rate are among the data acquired. This dataset is open to the public and comprises raw and highly processed data as well as extracted features. Using validated questionnaires relating to workload, psychological stress, and other characteristics, we evaluated work behavior and effect datasets.

III. METHODOLOGY

A. DEep Local Feature (DELF) Extraction and Dataset

The WESAD dataset was used in this research. Attila Reiss, Philip Schmidt, et al. introduced and made publicly available this dataset in 2018. Movement and physiological data were collected from 15 patients using The Empatica E4 arm sensor and the RespiBAN Advanced chest device. Subjects' physiological stimuli were recorded through several study protocols: preparation, baseline, fun state, stress state, meditation, and recovery. Details of the sensor setup, placement, and the method used to generate this data set are described in Ref. [19], including data collected during any patient study protocol. ACC, RESP, EMG, TEMP, and heartbeat were measured using RespiBAN.

All sensor signals were segmented using a sliding window with a 1-second shift. On the unprocessed ACC signal, various statistical parameters such as standard deviation, mean, min, max were independently calculated for the (x, y, z) axes and summed as absolute values for all axes (3D). Statistical characteristics such as mean, standard deviation, minimum and maximum values were calculated using raw ACC, ECG, RESP, and TEMP signals. also uses the TEMP peak frequency and signal slope δ_{TEMP} as characteristics. Statistical characteristics like standard deviation, mean, minimum, and peak values were obtained after filtering the unprocessed ECG signal through some kind of low pass filter with something like a 5 Hz [20].

By using the Fully Convolutional Networks (FCN) feature extraction approach, a dense grid of local features is produced. The result of a conv4 x convolution function within the ResNet50 model [21] is used to create FCN. An image pyramid is built, and then an FCN is applied to each level separately, to achieve scale invariance in discrete portions of the picture. The attributes that describe distinct image locations are represented by image pyramids. The functional map generates a dense grid of local features once FCN is performed to every level of the pyramid. Fine-tuning is done to improve the identification of local descriptors, making the method more robust. The ResNet50 model trains image classification using standard cross-entropy. 224X224 frames are randomly selected for training. Local descriptors are used in methods to examine image representations that can also be created using object or degree of patching labels.

By first conducting L2 normalization and then using PCA to decrease the dimensionality of the acquired feature descriptors to 40, The feature descriptors' dimensionality is decreased and filtered. The descriptors' 40 dimensions allow for good discrimination and compactness.

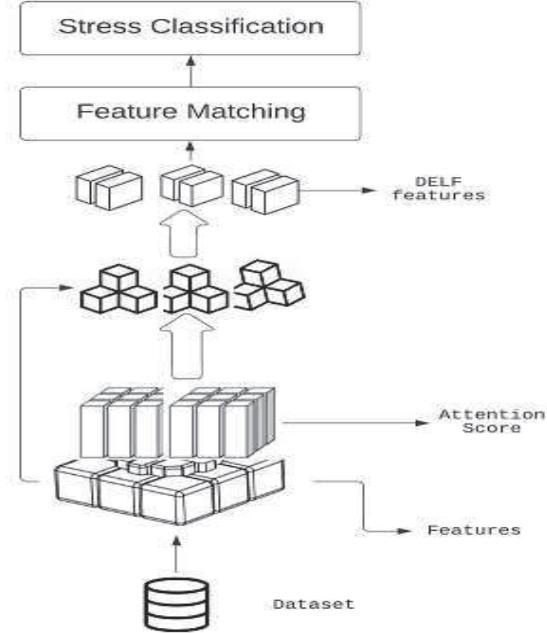


Fig:1- Proposed System Framework

The suggested system's results were compared to those of several other contemporary algorithms that produce local and global descriptors for a given image.

A total of four alternative algorithms are compared to the suggested system. Deep Image Retrieval (DIR) uses 2048-dimensional descriptors and gives multi-resolution descriptors. CONGAS [22] uses a Laplace or Gaussian cue point detector and then applies the Gabor wavelet response to the output of the cue point detector above to generate 40-dimensional feature descriptors. LIFT [23] harvests 128 special features while learning keypoint recognition, description, and orientation estimation. A new version of precision and recall is as follows for performance comparison:

$$P_{precision} = \frac{\sum_q |R_q^{TP}|}{\sum_q |R_q|} \quad R_{recall} = \sum_q |R_q^{TP}|$$

TP- True Positive

B. Feature Matching

The bagging/bootstrap aggregating strategy is used to learn from trees. The bootstrapping improves the model's performance since it reduces variance without raising bias. As a result, it is less susceptible to noise. The distinction between this approach and other tree learning algorithms is that it employs a random subspace method, in which a random subset of characteristics is employed at each candidate split. This randomization aids in the removal of any correlation between any two trees in the forest, which is a significant component impacting the forest error rate. Given m predicted features, each one of the s observations X_j is generated independently according to a weighted mixture density of the form $g(x) = \sum_{i=1}^m w_i * g_i(x - y_i)$. In this way, each observation is effectively associated with all of the predictions. Note that with this model, all observations might, with one particular realization, lie near a single y_t . The joint density function is thus:

$$f_{x/h}(X, S / H) = \prod_{x_j \in X} \left(\sum_{i=1}^m q_i * g(x_j - y_i * \theta_i) / \sum_{i=1}^m q_i \right)$$

the observations are divided in two disjoint subsets of matched features and spurious features and the resulting joint conditional density function can be written as

$$f_{x/H}(X, S / H) = P_s(s) * \prod_{x_j \in X} \left(\sum_{i=1}^m q_i * g(x_j - y_i * \theta_i) / \sum_{i=1}^m q_i \right) * \prod_{x_j \in X} \rho(x_j)$$

C. Stress Categorization

We divide the feature maps and labels into training and test sets after obtaining them. The majority of the data is utilized for learning (18,101,440 features) while the remaining 20% is used for testing. We utilize scikit Learn's "Train Test Split" preset function for this randomization. Data is fed into an existing model, which is used to train and categorize additional data. At this point, you must select a model, learn about its properties, and adjust its settings. We used DNN and LSTM to categorize the level of stress due to physiological data in our study.

D. Model's Evaluation

a. Deep Neural Network Model

To categorize stress, we initially evaluated a simple DNN architecture, as shown in table I.

TABLE I. THE DNN MODEL'S PARAMETERS

Characteristics	Defined Values
No.of.Layers	50
Dropout	0.25
Activation Function	comb-H-sine
Optimizer	Adam
Loss Function	Categorical Cross-Entropy

The class accuracy for the preset version configurations averaged 82%. Classification outcomes may be stepped forward with the aid of using making the version layers denser, growing the variety of epochs, growing the scale of hidden layers, and so on.

b. LSTM network model

LSTM is specifically designed to prevent long-term toxicity problems. The default behavior is to remember information for a long time. It determines whether data is saved, outputted, or forgotten using

memory cells and gates. The output of the bottom layer and the result of the preceding time unit of the current layer are sent into the LSTM. For this test, the Long short - term memory classification model was chosen and its parameters are listed in table2.

TABLE 2-THE LSTM MODEL'S PARAMETERS

Characteristics	Defined Values
No.of.Layers	50
Dropout	0.25
Activation Function	comb-H-sine
Optimizer	Adam
Loss Function	Binary Cross-Entropy

The accuracy of the classification is 86 percent on average.

c. Evaluation of the classification rate

We tested both Long short - term memory and DNN performances while increasing the number of epochs in the tests. The graph below depicts the progression of accuracy through time for various epochs.

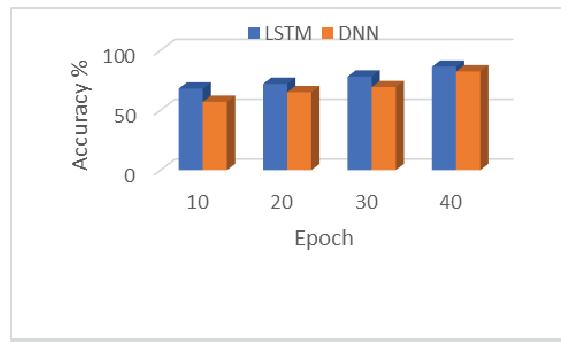


Fig2: - Classification rate Accuracy

When the number of epochs grows, there is a modest rise inaccuracy rate. At 30 and 40 epochs, we got the best accuracy rate (82%) using the DNN model. Increasing the number of epochs takes time and always results in a longer training period. We may claim that the optimum epoch number for this model that allows us to achieve the highest precision score is 30. The LSTM version outperformed the DNN version. The maximum

accuracy rating recorded is seventy-six percent, which becomes reached after 30 epochs. To summarize, each model achieved higher than expected and with the preselected parameters for every version, the LSTM outperformed the DNN significantly.

IV. RESULTS AND DISCUSSION

The proposed task defines two categorization tasks to identify stress according to a person's emotional state. To evaluate the model, we used Mean Absolute Error, Root Mean Square Error, Pearson correlations, and Spearman correlations that take into account the ordinal structure of the response variables. In our experiments, we used two classifiers: DNN and LSTM.

TABLE 3-PERFORMANCE COMPARISON

	DNN Model		LSTM Model	
	Non-stress	Stress	Non-stress	Stress
Sensitivity	0.57	0.34	0.82	0.59
Specificity	0.68	0.8	0.81	0.91
Precision	0.58	0.32	0.77	0.66
Recall	0.63	0.72	0.76	0.67

From the above table, we can conclude that the LSTM model gives better performance.

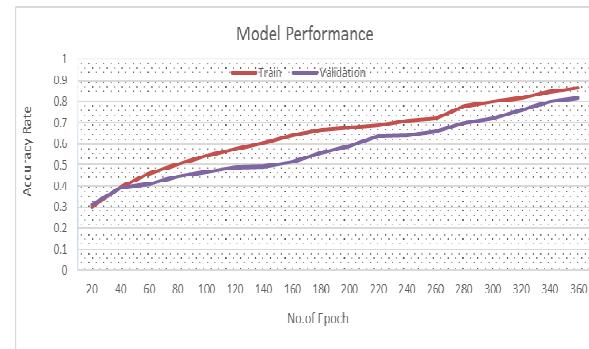


Fig 3-Stress Recognition Accuracy

The training and validation accuracy attained when training the model on the dataset are shown in Figure 3. The validation accuracy is 82 percent, whereas the training accuracy is 86 percent.

V. CONCLUSION

This study proposes a descriptor-based stress recognition application based on deep local features, developed from a CNN-based model. To properly assess the effectiveness of ensemble classifiers, we ran experiments on the WESAD dataset. LSTM and DNN are often employed for stress identification based on physiological sensors. Based on the results, we can conclude that the LSTM model is more efficient and accurate than the DNN in our scenario and given the parameters we have set. The quality of such sensing devices, the reliability of the samples, and also the number of features offered during learning, which is the largest issue, all have an impact on the model's performance. In the future, we will focus on identifying and integrating additional parameters while maintaining a non-invasive approach. This will allow more features and methods to be considered, which will increase the accuracy of the response. Pattern recognition and classification techniques will be used to detect emotions and work to understand the user's mood. These techniques, in addition to eye variables (eye gaze detection, blink frequency), will strengthen the conflict resolution process by complementing stress detection.

REFERENCES

- [1] J. S. Wang, C. W. Lin, and Y. T. C. Yang, "A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition," *Neurocomputing*, vol. 116, pp. 136–143, 2013.
- [2] J. Zhai and A. Barreto, "Stress detection in computer users through non-invasive monitoring of physiological signals," *Biomed. Sci. Instrum.*, vol. 42, pp. 495–500, 2006.
- [3] M. Singh and A. Bin Queyam, "Stress Detection in Automobile Drivers using Physiological Parameters: A Review," *Int. J. Electron. Eng.*, no. 2, pp. 1–5, 2013.
- [4] N. Kawakami, F. Kobayashi, S. Araki, T. Haratani, and H. Furui, "Assessment of job stress dimensions based on the job demand-control model of employees of telecommunication and electric power companies in Japan: reliability and validity of the Japanese version of the Job Content Questionnaire," *Int. J. Behav. Med.*, vol. 2, no. 4, pp. 358–375, 1995.
- [5] N. Mucci et al., "Work-related stress assessment in a population of Italian workers. The Stress Questionnaire," *Sci. Total Environ.*, vol. 502, pp. 673–679, 2015.
- [6] U. Bashir and M. Ismail Ramay, "Impact of stress on employees job performance: A study on banking sector of Pakistan," *Int. J. Market. Studies.*, vol. 2, no. 1, pp. 122–126, May 2, 2010.
- [7] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, "Machine learning framework for the detection of mental stress at multiple levels," *IEEE Access*, vol. 5, pp. 13545–13556, 2017, DOI: 10.1109/ACCESS.2017.2723622.
- [8] H. Kuper, M. Marmot, and H. Hemingway, "Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease," *Seminars Vasc. Med.*, vol. 2, no. 3, pp. 267–314, 2002, DOI: 10.1055/s-2002-35401.
- [9] S. Abdullah and T. Choudhury, "Sensing technologies for monitoring serious mental illnesses," *IEEE Multimedia*, vol. 25, no. 1, pp. 61–75, Jan.–Mar. 2018, DOI: 10.1109/MMUL.2018.011921236.
- [10] R. Cowie et al., "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.
- [11] R. W. Picard, "Affective computing: From laughter to IEEE," *IEEE Trans. effect. Comput.*, vol. 1, no. 1, pp. 11–17, Jan. 2010.
- [12] B. de Gelder, "Why bodies? Twelve reasons for including bodily expressions in affective neuroscience," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 364, no. 1535, pp. 3475–3484, Dec. 2009
- [13] P. C. Schmid, M. Schmid Mast, D. Bombari, F. W. Mast, and J. S. Lobmaier, "How mood states affect information processing during facial emotion recognition: An eye-tracking study," *Swiss J. Psychol.*, vol. 70, no. 4, pp. 223–231, Dec. 2011.
- [14] D. J. Martino, S. A. Strejilevich, G. Fassi, E. Marengo, and A. Igoa, "Theory of mind and facial emotion recognition in euthymic bipolar I and bipolar II disorders," *Psychiatry Res.*, vol. 189, no. 3, pp. 379–384, Oct. 2011.
- [15] G. Sandbach, S. Zafeiriou, M. Pantic, and L. Yin, "Static and dynamic 3D facial expression recognition: A comprehensive survey," *Image Vis. Comput.*, vol. 30, no. 10, pp. 683–697, Oct. 2012.
- [16] G. Giannakakisa, M. Pediaditisa, D. Manousos, "Stress and anxiety detection using facial cues from videos", Elsevier 2016.
- [17] Jacqueline Wijsman, Bernard Grundlehner, Hao Liu, "Towards Mental Stress Detection using wearable physiological sensors", IEEE 2011.
- [18] Saskia Koldijk, Maya Cappelli, Suzan Verberne, "The SWELL Knowledge Work dataset for stress and user modeling research", ICMI 2014.
- [19] Philip Schmidt, A. Reiss, R. Duerichen, Kristof Van Laerhoven, "Introducing WESAD, a multimodal dataset for wearable Stress and Affect Detection", International Conference on Multimodal Interaction 2018.
- [20] S. Kwon, "Att-Net: Enhanced emotion recognition system using the lightweight self-attention module," *Appl. Soft Comput.*, vol. 102, Apr. 2021, Art. no. 107101.
- [21] M. Raja and S. Sigg, "Applicability of RF-based methods for emotion recognition: A survey," in Proc. IEEE Int. Conf. Pervez. Comput. Commun. Workshops (PerCom Workshops), Mar. 2016, pp. 1–6.
- [22] N. Dael, M. Mortillaro, and K. R. Scherer, "Emotion expression in body action and posture," *Emotion*, vol. 12, no. 5, p. 1085, 2012.
- [23] H. Aviezer, Y. Trope, and A. Todorov, "Body cues, not facial expressions, discriminate between intense positive and negative emotions," *Science*, vol. 338, no. 6111, pp. 1225–1229, 2012.

Performance of Indian Cricket Team in Test Cricket: A comprehensive Data Science analysis

Vishnusai Viswajith Tharoor

Department of Computer Science and Engineering

Amrita School of Engineering, Coimbatore

Amrita Vishwa Vidyapeetham, India

cb.en.u4cse18068@cb.students.amrita.edu

Dr Dhanya N.M

Department of Computer Science and Engineering

Amrita School of Engineering, Coimbatore

Amrita Vishwa Vidyapeetham, India

nm_dhanya@cb.amrita.edu

Abstract—Cricket is a sport that is widely hailed in India. It is followed fervently in the country, such that some players are worshiped, and are considered holy. Created by shepherds of England, it has now grown significantly, and equipped some of the most advanced technologies to its armory, which help players improve their performances, and also plan strategies against opposition teams. The proposed work aims at analyzing the performance of the Indian Cricket Team in Test format. All oppositions are considered, and the data is analyzed to extract valuable information, which could help serve as an indicator for future performances. Exploratory Data analysis and Visualization was done using python libraries such as pandas, numpy, matplotlib, seaborn. Classification models were implemented for the data based on number of overs, and their results were analyzed. Random Forest Classifier was the most effective among the models implemented, with an accuracy of 75%.

Index Terms—ICT(Indian Cricket Team), Test Cricket, DataFrame, Exploratory Data Analysis(EDA), Machine Learning(ML), Classification.

I. INTRODUCTION

A game of cricket has two teams competing against each other. During the start of the game, a coin toss is done by one of the team captains, and the other team captain makes a call(heads/tails). The winner of the toss makes a decision to either bat or bowl. There are three main entities for a game. They are

- Batter
- Bowler
- Fielder

A batter scores runs, and a bowler takes wickets. The fielder aids the bowler in taking wickets, and also to ensure batters don't get easy runs. A win is determined by the aggregate number of runs scored by the team. An inning is the complete duration of one team's batting/bowling performance. At the end of the match, the team with the greater aggregate of runs, is determined as the winner. However, test match cricket is different. The overall game would comprise of minimum 3 innings, to a maximum of 4 innings in normal circumstances. A team has to play a minimum of one inning, and a maximum of two innings. The game lasts for a total duration of 5 days. Various factors like toss, pitch conditions, weather conditions play a role in determining the winner of the game. The following are the results possible for a given game:

- Win
- Draw (or) Tied
- Loss

The proposed work involves Exploratory Data Analysis on the data-set using python, and derives statistical inferences based on the analysis and visualization results obtained. It also deals with mining rules from the data-set. It also implements classification models on the data-set in order to understand the significance of attributes.

II. RELATED WORK

Various articles related to Test format of cricket were published before the advent of the millennium. However, all exploratory data analysis related to cricket has recently evolved. Comparison of ICT's performance in two periods was done using Association Rule Mining. The first was from 1974 to 2000. The next was from 2001 to 2010. The ICT's performance had substantially improved in the second period considered. [1].The Exploratory data analysis on IPL data , authored by Mohapatra et al., [2] focuses on performing data analysis for IPL data, as well as presenting a model that helps select the team, and also predict the winner for a given IPL game. Regression was mainly used in order to perform the task. It considers various factors of the game like venue, toss, team, track record, and team that is to play [2] .Pandas is the go-to library when it comes to performing data analysis using python. DataFrame is the data structure that is preferred in python. Libraries like seaborn, matplotlib help in visualizing the data and obtain useful results. Bar plots, scatter plots, box plots are some of the plots that could help obtain valuable insights on performances. [3] [6] [8].Classification could be done using supervised and unsupervised methods, depending on the data present. Support Vector Machines are accurate in predictions, and have the ability to handle high dimensions in data. Random Forest method takes in an ensemble of decision trees, and the final class is decided as the major result provided by a large subset of that ensemble. [8] [12]. Random forest method can be performed for both regression and classification. [4] Also, a simple means of classification using an artificial neural network is the Perceptron. It helps identify the linear separability of data as well. [6] [12].

match_id	season	start_date	venue	innings	ball	batting_team	bowling_team	striker	non_striker	...
2603	291352	2007/08	2008-01-02	Sydney Cricket Ground	4	70.2	India	Australia	R.P. Singh	A.Kumble
2604	291352	2007/08	2008-01-02	Sydney Cricket Ground	4	70.3	India	Australia	I.Sharma	A.Kumble
2605	291352	2007/08	2008-01-02	Sydney Cricket Ground	4	70.4	India	Australia	I.Sharma	A.Kumble
2606	291352	2007/08	2008-01-02	Sydney Cricket Ground	4	70.5	India	Australia	I.Sharma	A.Kumble

Fig. 1. Sample data - Australia vs India, Sydney-2008

III. DATA SET

Data set is obtained from cricsheet.org . Cricsheet is a collection of projects which collectively provide data for various aspects of cricket. Ball-by-ball raw data folder, pertaining to all test matches from 2004 till January 11,2022 is chosen. The main data set contains information about 627 test matches that have occurred in this period. The raw data has important attributes pertaining to a given delivery, such as start date, batter, bowler, batting team , bowling team no. of overs, run scored of that delivery, whether it resulted in a wicket, and so on. The raw data is available as two types:

- CSV files that comprises ball-by-ball data
- JSON files that comprises both ball by ball data, as well as off-the-field information such as venue, toss winner, winning team, player-of-the-match etc.

ICT has played close to 178 test matches in this period, and all these information are taken for further analysis. Fig. 1 is an illustration of the raw data for a given match

IV. EXPLORATORY DATA ANALYSIS

EDA involves understanding the type of data, their statistical distributions, handling missing data, and analyzing and understanding the data as a whole. [5].

A. Data pre-processing

The raw data has a detailed description of every delivery, for a given match. In order to obtain valuable insights about ICT's performance, pre-processing is necessary. So, for a given test match, off-the-field information such as year, venue, opponent, toss winner, India's role in the first innings, winner, player-of-the-match, were obtained from the JSON file pertaining to that particular match, which was identified using the match-id attribute in Figure 1. The same attribute was used to compute the inning-wise score, wickets taken , and top performance based on runs or wickets from each team. The innings score is computed, when the data is grouped by innings, and the sum of runs scored in that particular inning is computed. A similar approach is followed to compute other continuous attributes. Another attribute present is the Captain for the game. Two captains are considered: MS Dhoni and Virat Kohli. Although other players have captained the team for a few matches, ICT was following major trends during their tenure. This attribute could really help indicate the nature of performances given by the team. Fig. 2 denotes the pre-processed data obtained.

Year	Venue	Opponent	Choice	Captain	1st innings score	1st innings no. of overs	1st innings wkt's lost	1st innings top score	1st innings top wickets	...	Toss_Winner	Winner	MoM
173	2021	Kemington Oval, London	England	Batting	V.Kohli	191	61.3	10	57	4 ...	England	India	RG Sharma
174	2021	The Rose Bowl, Southampton	New Zealand	Batting	V.Kohli	217	92.1	10	49	5 ...	New Zealand	New Zealand	KA Jamieson
175	2021	Wankhede, Mumbai	New Zealand	Batting	V.Kohli	325	109.5	10	150	10 ...	India	India	MA Agarwal
176	2021	SuperSport Park, Centurion	South Africa	Batting	V.Kohli	327	105.3	10	123	6 ...	India	India	KL Rahul
177	2022	The Wanderers Stadium, Johannesburg	South Africa	Batting	V.Kohli	202	63.1	10	50	4 ...	India	South Africa	D Elgar

Fig. 2. Sample pre-processed data of ICT

B. Data cleaning

After the following data was obtained, data cleaning was performed. On analysis, it was found that the 4th innings mostly had missing values. The reason behind this, is that Test cricket allowed for wins that involved an innnings' defeat. In this case, the winning team doesn't perform an extra inning, as the one inning performed has earned them a win. In some cases matches were drawn due to unfavorable weather or unforeseeable circumstances. Filling the missing value with 0 would necessarily mean that the analysis of the attributes could lead to incorrect findings. Hence, the missing values were replaced with median, as the median is sensitive to outliers, and it also doesn't impact the other values present.

C. Statistical insights

During this period, it was found that the average total by ICT in the 1st innings, is 367, and in the 2nd innings, is 386, and in the 3rd innings ,is 252, and in the 4th innings, is 172. ICT has won 49 games batting first, and 38 games chasing. This indicates that the team has been more successful batting first. Also, the team has drawn 26 games while chasing, and 21 games while batting first. This indicates that it could also prevent a loss by batting sensibly in the last innnings.The most successful chase by the team, in this period, has been 387. The highest total scored by the team, in this period, is 759. The lowest score by the team for any innnings, is 36. In terms of bowling, the best bowler in the team would bag a minimum of 4 wickets, when it came to wins. The no. of times the bowler has taken 10 wickets, is comparatively more in Virat Kohli's leadership, than M.S Dhoni's leadership.

1) *Batters statistics:* Fig.3 denotes the data derived from raw data about batters. The raw data is grouped by innnings, and if the batting team is India, then the details about batters are stored and later converted into a pandas DataFrame. Kohli is the most successful batter during this period. In a span of 11 years, he has scored 7854 runs, with an average of over 50. Conversion rate is the percentage of 50 plus scores that have been converted into hundreds. Kohli has the best conversion rate among batters. VVS Laxman has hit 35 half centuries, the most during the period. MS Dhoni has the worst conversion rate among batters who have played more than 100 innnings. I Sharma has scored the most ducks or zeroes during this period. Strike rate does not play a major role in test cricket. [16].

2) *Bowlers statistics:* Fig. 4 denotes the data derived from raw data about bowlers. The raw data is grouped by innnings, and if the bowling team is India, then the details about bowlers are stored and later converted into a pandas DataFrame.

Player	Innings	Runs	Average	HS	Ducks	200+	100s	50s	Conversion rate
V Kohli	166	7854	51.31	254	14	7	27	27	50.00
CA Pujara	160	6661	45.63	206	11	3	18	32	36.00
SR Tendulkar	133	5969	48.88	248	2	3	17	28	37.78
R Dravid	134	5708	46.60	270	3	1	16	26	38.10
V Sehwag	125	5525	48.20	319	13	3	14	23	37.84

Fig. 3. Best batters of ICT in tests (2004-present)

Ashwin is the most successful bowler in this period. Even though he made his debut in 2011, he has taken the most wickets during this period. The number of five wicket hauls he has taken, is the highest amongst Indian bowlers, by a huge margin. He is also the player who has played the most number of innings as well.

3) *Captaincy statistics:* The data is derived by filtering matches based on wins, draws and losses. Win loss ratio, and win percentage are derived attributes. MS Dhoni and V Kohli have contrasting captaincy styles, and it is reflected in the match results as well. Under MS Dhoni, India had won 31 test matches, and lost 24, which is contrasting the captaincy tenure of V Kohli. India had won 44 games, and lost just 16. In the literacy survey, it was already found that India was better performing in the 21st century, compared to its previous years. This effectively proves that V Kohli is the most successful captain for India in the test format. It is also notable that Virat Kohli has won 7 player-of-the-match awards as captain. Usually, captaincy tends to affect the bowling/batting performance of the player. V Kohli has been the best batter, as well as the best captain for the ICT in the test format, during this period.

4) *Nation wise statistics:* The data is derived, by grouping the pre-processed test matches data based on opposition, and then aggregating the wins, draws and losses based on Dataframe filtering. In this period, ICT has the most wins against Australia and England. It is also notable that in this period, it has played the most number of matches, as well as lost the most number of matches against Australia and England. India has been most successful against the West Indies, as it has played 20 games so far, and hasn't lost a single test match. It can also be concluded that England has been the most difficult team to beat in this period. Overall, India has fared well against all oppositions, as there is no scenario where the number of losses exceed the number of wins.

D. Data Visualization

Data Visualization is a measure that would help users gain insights on data through various graphical plots and illustrations. Cricket analysts use data visualization tools in order to show various useful relations between various data attributes.

Fig. 6 depict ICT's annual aggregate wins in the considered period. This data is derived by filtering the data based on wins achieved by ICT in matches, and aggregating the total

Player	Innings	Wickets	5WH
R Ashwin	140	450	34
I Sharma	138	317	11
RA Jadeja	89	238	9
Harbhajan Singh	87	236	12
Z Khan	90	220	8

Fig. 4. Best bowlers of ICT in tests (2004-present)

Opponent	Total matches	Wins	Draws	Losses
Australia	37	16	10	11
England	40	16	9	15
South Africa	27	13	5	9
West Indies	20	12	8	0
Sri Lanka	19	11	4	4
New Zealand	18	8	6	4
Bangladesh	9	7	2	0
Pakistan	6	2	3	1
Zimbabwe	1	1	0	0
Afghanistan	1	1	0	0

Fig. 5. Performance of ICT against each nation

number for each year. 2016 is the most successful year for ICT . With 9 wins in the test matches held. 2010, 2021 are the second most successful years for Indian Cricket. Additionally, the same steps were repeated to visualize India's toss record, as shown in Fig. 7. India won the most tosses in the year 2008.

The annual number of centuries were derived as a DataFrame from the raw data. Each game in the dataset was grouped by the innings that was being played. Then, the data was filtered based on ICT being the batting team, and the overall sum for each player was aggregated. Year, No. of Centuries are the data involved. 2010 was the most successful year for ICT in terms of the no. of centuries. 23 centuries were hit in that year, signifying the batting was at its peak during that year. 2016 and 2017 saw the most no. of centuries that resulted in wins. Fig. 9 depicts the centuries hit by ICT players.

Fig. 10 denotes that more the number of innings played, more the probability for a batter to score more runs. The trend followed so far, is that if a player exceeds more than 120 innings, the increase in the no. of runs seemed to be following an exponential trend.

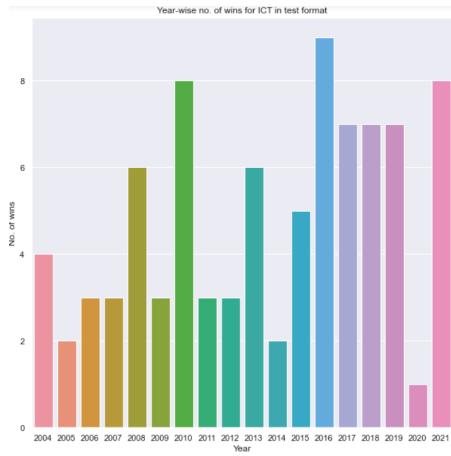


Fig. 6. ICT - Aggregate annual wins

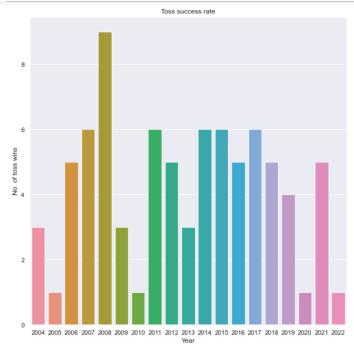


Fig. 7. ICT - Aggregate toss wins

Fig. 8 is a box plot between runs scored in the first innings for each year, based on choice. Blue plots denote the runs conceded, and red plots denote the runs scored. Until 2014, the blue plots mostly were plotted above the reds, denoting that the runs conceded were higher than runs scored. Since 2015, the red plots were consistently plotted higher than the blue plots. Until 2014, MS Dhoni was the Indian Captain. After 2014, V Kohli was the Indian captain. This indicated the improvement of tactics against opponents, and also the improvement in bowling performances.

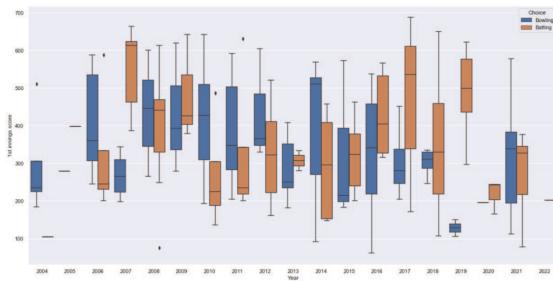


Fig. 8. Box plot between runs scored in 1st innings each year based on choice

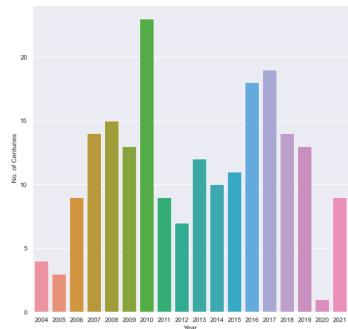


Fig. 9. ICT - Aggregate annual centuries

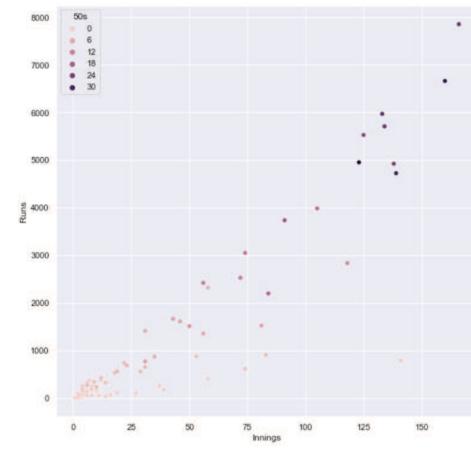


Fig. 10. Scatter plot between innings and runs scored- ICT



Fig. 11. Heat map of confusion matrix between toss wins and match results

Fig. 11 denotes the heat map of a confusion matrix between toss wins and match results. This was done in order to check the impact of toss result on the overall match results. 55% of the games played, had teams that won the toss, won the match. Unlike one day games, toss plays a vital role in test format, as there may be scenarios where the team that wins the toss would be choosing a decision that would provide suitable

conditions, and thus create winning situations.

Additionally, a bar chart was created to visualize the aggregate player-of-the-match awards given to each player, and from the plot, it was inferred that V Kohli and R Ashwin were awarded the most awards. Both players got 7 awards during this period. It is also notable that V Kohli and R Ashwin are the most successful among batters and bowlers respectively. Thus, the better an individual player performs, the more awards that could be won.

V. ADDITIONAL PRACTICES

A. Association rule mining - Apriori Algorithm

Apriori algorithm is used to find out association rules between objects [1]. Association rules are derived from frequent item-sets, which is a resultant of the algorithm [15]. For the considered data, attributes were considered as objects, and associations between attributes was to be found. The ICT Test matches data was first split into two data sets:

- Data set where ICT batted first
- Data set where ICT bowled first

The first data set was grouped by the 1st innings score, and second was grouped by the 2nd innings score. This was done in order to emphasize on the batting performance on both the innings. After encoding, and ensuring that the data is in the right format, the model is implemented with all proper hyperparameters, and the results are given in Fig. 12 and Fig. 13.

When a deeper understanding of these results are made, it can be understood that when a team plays more no. of overs, the magnitude of the top scorer of the innings would be high. Subsequently, the innings score would be high. This could help set up a good 3rd innings performance and subsequently win the game.

When the team bowls first, the lower the top scorer, the better the bowlers have been able to capitalize on taking wickets. If the 1st innings score is less, then it encourages the team to perform better in the second innings. If a significant score is posted, the team chasing could be easily favourable for a win.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
217	(1st innings no of overs)	(3rd innings score, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
438	(1st innings no of overs, 1st innings score)	(3rd innings score, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
439	(1st innings no of overs)	(1st innings score, 3rd innings score, 1st inn...	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
543	(1st innings no of overs, Year)	(3rd innings score, 1st innings top score, 2nd inn...	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf
548	(1st innings no of overs)	(3rd innings score, Year, 1st innings top score)	0.0375	0.075	0.0375	1.0	13.333333	0.034687	inf

Fig. 12. Apriori algorithm results for batting first

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
299	(1st innings top score)	(2nd innings score, 1st innings score, 3rd inn...	0.049383	0.481481	0.049383	1.0	2.078923	0.025606	inf
313	(2nd innings no of overs)	(2nd innings score, 1st innings score, 3rd inn...	0.037037	0.481481	0.037037	1.0	2.078923	0.019204	inf
404	(Year, 1st innings top score)	(3rd innings score, 1st innings score, 2nd inn...	0.049383	0.481481	0.049383	1.0	2.078923	0.025606	inf
412	(1st innings top score)	(3rd innings score, 1st innings score, 2nd inn...	0.049383	0.481481	0.049383	1.0	2.078923	0.025606	inf
433	(Year, 2nd innings no of overs)	(3rd innings score, 1st innings score, 2nd inn...	0.037037	0.481481	0.037037	1.0	2.078923	0.019204	inf

Fig. 13. Apriori algorithm results for bowling first

TABLE I
CLASSIFICATION RESULTS

Classifier model	Accuracy
Random forest classifier	75%
Support Vector Machine	58%

B. Hypothesis Testing

Hypothesis testing is the method used to check if a statistical statement can be accepted or rejected, by means of certain computations [14]. The following tests were done:

- Mean of number of wins each year is greater than 7
- Mean of number of Centuries after 2014 is lesser than 7

These tests are one tailed [14]. Significance value for the hypotheses considered are 0.05 and 0.1 respectively. Both hypotheses were accepted. The first hypothesis is an indicator of how the team has fared better under V Kohli's captaincy. The second hypothesis is an indicator of how the batters are relatively making lesser impact, and how bowlers have gained prominence.

C. Classifier models

Based on the data input, a classifier categorizes it into different classes. In the proposed work, models have been created based on the number of overs. There are two classes: 0 denoting ICT hasn't won the particular game considered; and 1 denoting ICT winning the game. Random Forest Classifier and Support Vector machine were the two models chosen [7]. Random Forest Classifier helps improve the accuracy of the results, due to the consideration of multiple decisions, and Support Vector Machines work very efficiently in linearly separable data. Table I denotes the results obtained. Random Forest Classifiers provided 75% accuracy. The reason for Support vector machines not producing favourable results, may be due to the fact that the data is not linearly separable. It also denotes the fact that the number of overs being played in a test match involving ICT has a role to play in deciding the winner.

VI. CONCLUSION

A. Findings

On performing the study, it can be said that since 2004, there has been an upward progression in terms of the no. of wins, and also in the overall team performance. This has been concluded using the win-loss ratio amongst opponents as the metric to measure the overall success. Various notable players have retired from the game, but it has not affected the overall performance of the team. Since 2014, there has been an upward trend in the overall performance of the team. The number of losses during that period were relatively low, to the losses before this period. Win-loss ratio and ratio between wins and sum of draws and losses, are the metrics used to measure the abilities of captains. It can be inferred through the analysis, that V Kohli has been the best captain in this period. A batter's performance is best measured by the conversion rate. The findings suggest that V Kohli has been the best performer

of the team. Apart from these findings, it can be noted that the overall performance of batters after 2018 has been low, and it needs to be improved. Amongst the classifier models tested, Random Forest Classifier has the best accuracy relatively, to showcase the effect of the no. of overs on the match result.

B. Future enhancements

The data that could be extracted on bowlers were relatively low. More information could be extracted, and analysed. Home-Away match performance could not be interpreted, and based on Geo-Encoding, accurate locations could be extracted, and performance at home, and away, could be analysed. When subsequent matches are played, more data would be thus present. The pandas library has been the go-to solution to handle table-like data in python. Terality is a newly developed library, that has the same syntax as python, but has a relatively better performance in all aspects [10]. This library could be used in order to perform operations. More models, such as deep learning models could be used in order to produce more insights on the data. [11] [16].

REFERENCES

- [1] Raj, K. A., Padma, P. (2013). Application of association rule mining: A case study on team india. 2013 International Conference on Computer Communication and Informatics. <https://doi.org/10.1109/iccci.2013.6466294>
- [2] Mohapatra, S., Goswami, A., Singh, A., Singh, V. K., Sen, B., & Sharma, K. (2021). Exploratory Data Analysis on IPL data. Lecture Notes in Networks and Systems, 315–325. https://doi.org/10.1007/978-981-16-4244-9_26
- [3] Pani, D. S. (2019). Exploratory data analysis using Python. International Journal of Innovative Technology and Exploring Engineering, 8(12), 4727–4735. <https://doi.org/10.35940/ijitee.l3591.1081219>
- [4] Raundale, P., Thosar, C., & Rane, S. (2021). Prediction of parkinson's disease and severity of the disease using machine learning and deep learning algorithm. 2021 2nd International Conference for Emerging Technology (INCET). <https://doi.org/10.1109/incet51464.2021.9456292>
- [5] Saranya, E., & Sivakumar, P. B. (2020). Data-driven prognostics for run-to-failure data employing machine learning models. 2020 International Conference on Inventive Computation Technologies (ICICT). <https://doi.org/10.1109/icict48043.2020.9112411>
- [6] M.S Suchitra, & L Pai, Maya (2017). Machine Learning using Advanced Python. <http://www.csi-india.org/>
- [7] Sivarajani, S., Ananya, S., Aravindh, J., Karthika, R. (2021). Diabetes prediction using machine learning algorithms with feature selection and dimensionality reduction. 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). <https://doi.org/10.1109/icaccs51430.2021.9441935>
- [8] Pisner, D. A., & Schnyer, D. M. (2020). Support Vector Machine. Machine Learning, 101–121. <https://doi.org/10.1016/b978-0-12-815739-8.00006-7>
- [9] Pajankar, A. (2022). Hands-on Matplotlib. <https://doi.org/10.1007/978-1-4842-7410-1>
- [10] <https://towardsdatascience.com/good-bye-pandas-meet-terality-its-evil-twin-with-identical-syntax-455b42f33a6d>
- [11] <https://www.geeksforgeeks.org/ipl-score-prediction-using-deep-learning/>
- [12] Dogru, N., & Subasi, A. (2018). Traffic accident detection using random forest classifier. 2018 15th Learning and Technology Conference (L&T). <https://doi.org/10.1109/ltn.2018.8368509>
- [13] <https://www.geeksforgeeks.org/implementing-apriori-algorithm-in-python/>
- [14] <https://towardsdatascience.com/hypothesis-testing-in-machine-learning-using-python-a0dc89e169ce>
- [15] <https://www.geeksforgeeks.org/apriori-algorithm/>
- [16] Sarkar S, Mishra S, Kumar S. Development of a Comprehensive Multi-Factor Method for Comparing Batting Performances in One-Day International Cricket. IIM Kozhikode Society & Management Review. 2022;11(1):92-108. doi:10.1177/22779752211017260

Deep Learning based Classification of Cervical Cancer using Transfer Learning

1st Hemalatha K

Department of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
hemshema5@gmail.com

2nd V Vetriselvi

Department of Computer Science and Engineering
College of Engineering Guindy, Anna University
Chennai, India
kalvivetri@gmail.com

Abstract—Cervical cancer is responsible for 90% of all deaths in low- and middle-income nations (LMIC). Cervical cancer is still one of the most common cancers in women's that develops in the cervix lower part of the womb that connects to the vagina canal. Cervical cancers are caused by continual infection on their cervix with one of the human papilloma viruses (HPVs). The most commonly used screening test for early detection of abnormal cells and cancer is the Pap smear test. Manual screening, on the other hand, leads to human errors. It is possible to save lives by detecting cancer early and accurately. Transfer learning has made significant progress in the field of machine learning in recent years, and the use of transfer learning technology to cervical cancer image classification has emerged as a new experimental domain. This work presents a study of transfer learning frameworks InceptionResNetV2, VGG19, DenseNet201 and Xception networks pre-trained on ImageNet, to classify cervical images using the SIPaKMeD dataset.

Index Terms—Cervical cancer, deep learning, image Classification, inceptionresnetv2, vgg19, densenet201, xception.

I. INTRODUCTION

Cervical cancer is the fourth most prevalent disease worldwide (with an estimated 570000 new cases per year) and the second most common cancer in women, according to World Cancer Statistics [1]. Around 80% of the 500000 cases and 90% of the 250000 deaths each year occur in low- and middle-income countries when preventative measures are insufficient. The number of new cases is projected to increase in the future years as the world population grows and ages. Cervical cancer is caused by a long-term infection with HPV, according to research. Cervical cancer, on the other hand, is the most curable cancer if diagnosed early and properly treated. As a result, making a precise diagnosis for cervical cancer in its early stages is essential. The Pap smear test, also known as the Papanicolaou smear, is the most extensively used procedure in cervical cytology for the detection of atypical lesion and cervical cancer. Cervical cytology evaluation, on the other hand, requires the use of professional physicians, which is both expensive and time-consuming. As a result, researchers have been working to create a variety of automated Computer-Aided Diagnose (CAD) methods is used for quick, sensitive, and accurate detection for cervical cancer, which helps pathologists and doctors to diagnose and prevent cancer more effectively [14].

Deep learning is a popular AI approach for developing decision-support systems for medical image classification. End-to-end classification with deep learning models, on the other hand, necessitates a substantial amount of training data, which would be typically unavailable in the medical field. Transfer learning is one solution to this problem, in which a model learned on a large dataset being re-used (after re-training) in the current circumstance with a small dataset [15].

In our proposed framework we use pre-trained deep learning models that have been trained on ImageNet dataset a non-medical image data repository with over 1.2 million scenery images, and fine-tuned on the cervical image dataset, which eliminates the need for a huge number of datasets and the challenges related with multi-class classification with unequal data distributions. The SIPaKMeD dataset, which contains single-cell cervical cytopathology images, is used for testing our proposed method. Four pre-trained convolutional neural network architectures are used in this study. The InceptionResNetV2, VGG19, DenseNet201 and Xception networks, each with its own set of capabilities, were used to classify cervical cancer images and compare the performance of the four pre-trained networks.

II. RELATED WORK

In the last five years, deep learning (DL) in medical imaging has been a major research topic, bringing lots of new review publications. Asiedu et al.[1] described two methods: (1) methods for collecting features/diagnosis of various contrasts in cervical films, and (2) methods for automatic features extraction for acetic acid/Lugol's-iodine cervical films. In order to extract simple but effective colour and textural features from cervigrams labeled with pathology algorithms were developed. The framework achieved 81.3 sensitivity, 78.5% specificity and 80% accuracy. Sari et al.[7] proposed a unsupervised feature extraction for effective representation and classification of image data. Salient regions in an image were better identified, and a novel deep learning-based technique was that quantified the regions by trying to extract a set of features. Song et al.[8] presents a adaptive shape priors method which was collected from the cytoplasm's contour fragments and shape statistics to segment the overlapping cytoplasm of cells in pap smear images. The approach consistently outperformed current best

practices which uses overlapping cells. Kurnianingsih et al.[5] uses a Mask R-CNN (Mask Regional Convolutional Neural Network) and a relatively small Visual Geometry Group-like Network for whole cervical cell segmentation to classify the VGG-like Net which yields an accuracy score of more than 96%.

Xue et al.[12] outlines an effort to build a cervical cancer prediction model that uses deep learning and transfer learning techniques to recognize and categories cervix images. ConvNet, which will classify the cervix images, is created using the three models: InceptionV3, ResNet50, and VGG19. The results of the experiment showed that the InceptionV3 model outperformed Vgg19 and ResNet50 on the cervical cancer dataset. Yu et al.[18] experimented with a number of convolutional designs before settling on residual neural networks with batch normalisation and dropout. For each class, the loss was estimated using the multi-class method loss with a logarithmic scale. The author also used 4633 extra images for training and employed a variety of data sources due to the short size of the dataset strategies for augmentation.

Hu et al.[4] presents a method to create "deep learning"-based visual evaluation algorithm that could detect cervical cancer and precancer automatically. The deep learning-based method was trained and validated using digitised cervical images from screening captured with a fixed-focus camera. On a two-sided basis, all statistical tests were performed. The findings support the idea of using modern digital cameras to do automated visual evaluations of cervical images.

III. SYSTEM ARCHITECTURE

The research's goal is to perform classification on cervical images using deep learning technology. In cervical cells, precise classification of the cytoplasm and nuclei is a challenging task. Modern machine learning and deep learning technologies have seen a lot of success in the medical health field in recent years. Figure 1 depicts the detailed architecture of the earlier levels of every CNN model. Conventional deep learning architectures rely on convolutional neural networks (CNN). CNNs have been found to be tolerant of image noise and invariant to translation, rotation, and size transformations in a wide range of studies. The CNN design is made up of convolution, pooling, and fully connected layers. The main building element of CNN structure is the convolution layer, which extracts low- and high-level properties of a given image. Convolved features are reduced in size by average-pooling following the convolution layer [16]. Training deep CNNs from scratch using random weight initializations can take more time, and with medical data scarcity, the process becomes even more time-consuming and tedious.

However, because of the requirement of having a large dataset and tuning an equal number of parameters, CNN has considerable limitations. To construct a general-purpose system that can identify diverse types of image datasets, a combination of CNN and transfer learning is used. Transfer learning, in which the weights of the pre-trained model are

explicitly transferred to the new task, is a feasible alternative. Therefore, we systematically investigate the performance of InceptionResNetV2, DenseNet201, VGG19, and Xception openly available CNN on the SIPaKMeD dataset.

A. Data-Preprocessing

In preprocessing step the image is resized to alter the input images to the model's requirements. To improve the model's generalizability, which will allow it to handle larger datasets and improve classification accuracy while avoiding overfitting, The data augmentation approach is used to enlarge the size of the training data artificially because it functions as a regularizer and helps avoid overfitting while training the deep learning model.

- **Resizing:** The SIPaKMeD dataset we used to test the performances of the proposed technique are in Bitmap and range in size from (71 x 59) to (490 x 474) pixels. For the four CNN networks, the object is resized for DenseNet201 and VGG19 (224 x 224 pixels) and InceptionResNetV2 and Xception (299 x 299) pixels. We used Keras' "preprocess-input" function in this case, which modifies input images to meet the model's needs.
- **Data Augmentation:** The Keras "Image Data Generator" API is used to improve model performance. During the training period, the images are randomly modified. As a result, in each epoch the network analyses unlike samples, extending the model's generalizability. We've set the feature wise centre to false and the rotation range to 40 degrees in this process. The width shift range, height shift range, shear range, and zoom range have all been set to 0.2, the horizontal flip to True, and the fill mode to nearest.

After data augmentation, the enhanced images are fed into a pre-trained CNN, such as the InceptionResNetV2, DenseNet201, VGG19, and Xception, which is used to transmit learnt parameters for classification. Transfer learning which makes it easier by using a part of a model that has already been pre-trained on the ImageNet dataset. The earlier levels of every CNN model are frozen and these layers are responsible for capturing more generic features. The last layers of the network were then fine-tuned using a dataset of cervical cancer cells, which was used to train the network on more dataset-specific properties. Finally, the fully connected classifier has been fine-tuned. The hyper - parameters of the proposed method are fine-tuned using the images. In this study, we used four ConvNet models; it's a deeper neural network because of the thousands of hidden layers. The input layer receives the raw pixel values of an image, and the output layer generates neurons based on the number of output layers that the image contains. The class to which the input image fits can be determined from the generated output.

The Google research team created the Inception-ResnetV2 network, the pretrained CNN model is employed to initialize the first few layers for the fine-tuned network. It takes a 299 x 299 x 3 pixel input image and generates feature maps of various dimensions in several layers. These features are

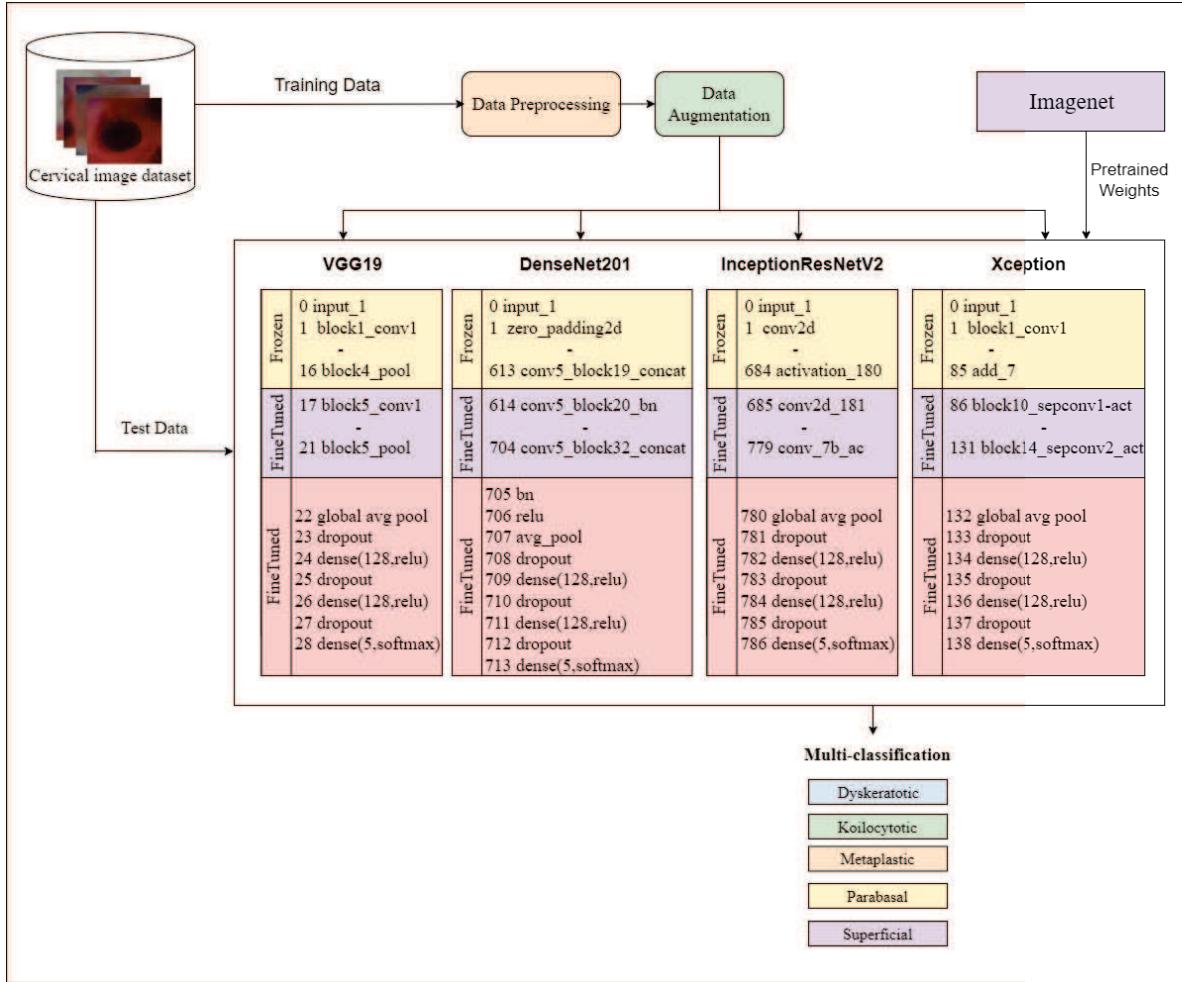


Fig. 1. Proposed Architecture

concatenated and given to next layer for deeper feature extraction. When the validation loss does not achieve a minimum during 20 consecutive epochs, early stopping is used to end the process. The VGG19 is a CNN created by the Visual Geometry Group at Oxford's and hence, the name VGG. The VGG19 is a modification of VGG models trained over the ImageNet dataset and contains 19 deep layers with an input image size of 244x244. In the VGG19, the kernel size is 3 x 3 with 1 stride size, while max-pooling is done in a 2 x 2-pixel window with a stride size of 2. There are different variants of the VGG such as VGG16 and many more. The main disadvantage of this CNN is the huge number of parameters that must be learnt. A pre-trained VGG19 model is used to extract and classify the features. In the first phase, only the fully connected layers are trained. The DCNN is trained on the top layer in the second stage, but the weights of some of the pre-trained network layers are fine-tuned.

In deep neural networks, DenseNet has been offered as a solution to the gradient vanishing problem. Dense blocks are

used to connect each layer directly to the next one, allowing for the creation of a dense network. To maintain the feed-forward nature, each layer gets additional inputs from all previous layers and passes on its own feature maps to all subsequent layers. This network accepts 224 x 224 x 3 pixels as an input image. DenseNets are made up of building blocks that are densely connected. As a result, the network will only need to learn a small number of parameters. Because fewer channels are handled in the convolutional layers, the number of training data is reduced, and the feature maps from previous layers are concatenated, the model becomes more compact. Xception is a type of convolutional neural network that employs Depthwise Separable Convolutions. A group of Google researchers developed it. This network accepts 224 x 224 x 3 pixels as an input image. We added fully - connected layers (2048) and a 5-way output units with Softmax activation function after the global average polling layer. It has been established that using the transfer learning strategy to train CNN with full training over the histopathology imaging modalities yields excellent

results, even when the training dataset is small. The obtained results and discussions are detailed in the next section.

IV. EXPERIMENTS AND RESULT ANALYSIS

In image classification, CNN models have been quite successful. The deep CNN model was trained entirely on a large image dataset (ImageNet), allowing it to pick up on many of the features needed for image classification. As a result, we made use of the concept of transfer learning for cervical cancer image classification in order to fully leverage the vast quantity of knowledge gained through pre-training the CNN model. There have been various CNN baseline models successfully applied to specific applications for image classification. InceptionResNetV2, DenseNet201, VGG19, and Xception were used to perform image classification of cervical cancer.

In this experiment, we used Google Colaboratory, a cloud service based on Jupyter notebook, to train and test our model. Many other machine learning libraries are pre-configured in Jupyter notebook for Python 2 and 3. These include Tensorflow, Matplotlib, Keras, PyTorch, and OpenCV. It allows users to practise DL in a Colab environment with a fully functional GPU (NVIDIA Tesla K80). The hyperparameter configurations for the experiment are listed in Table 1. We also opted with the Adam optimizer, a stochastic gradient descent variation with momentum and a learning rate of 0.0001. ReLU activations were employed to fine-tune the layers, followed by drop out 0.5. To quantify validation loss and accuracy, we employed network dropout. Because many of the connections were turned off, the network would appear to be underperforming. As a result, before loading our data into the model, we normalised and scaled each model's input images. The model was then trained for 100 epochs using the batches in our dataset, then the data was transmitted through the network to calculate the losses as shown in equation 1, the gradients were obtained, and the optimizer was run. We tested this model with a training and test dataset split 80/20.

$$Loss = - \sum_{i=1}^{outputsize} y_i \log \hat{y}_i \quad (1)$$

TABLE I
PARAMETER SETTINGS FOR THE FOUR TRANSFER LEARNING NETWORKS

Hyperparameter	Values
Initial Learning Rate	0.0001
Activation function	ReLU
Batch Size	32
Optimizer	Adam
Number of Epochs	100
Loss Function	categorical crossentropy
Momentum	0.9

A. Image Dataset

The SIPaKMeD dataset contains 4049 isolated cell images that were cropped separately using 966 cluster cell images

of Pap smear samples for the SIPaKMeD dataset. Expert cytopathologists divide the cells into five groups [3]. Normal cells are separated into two groups (superficial-intermediate and parabasal), aberrant but non-cancer are divided into two groups (koilocytes and dyskeratotic), and benign (metaplastic) cells are the final category. Sample images are shown in figure 2. For this study, the Single Cell Images (SCI) was employed individually [14]. Table II depicts the distribution of images in the collection. In order to evaluate the feasibility of the four

TABLE II
DISTRIBUTION OF SIPAKMED IMAGES IN THE COLLECTION

Class	Category	SCI
1	Superficial- Intermediate	831
2	Parabasal	787
3	Koilocytotic	825
4	Metaplastic	793
5	Dyskeratotic	813
-	Total	4049

CNN models in classifying cervical cancer accuracy, precision, recall, and F1-score are the most commonly used evaluation metrics. As a result of these factors, the F1-score is a measure that takes into account the accuracy and precision of the classifier, as well as how many samples are correctly classified. In this study, the samples of the classifications researched at this time are positive samples, whereas the samples of other classifications are negative samples.

For cervical cancer SIPaKMeD images we used the InceptionResNetV2, VGG19, DenseNet201 and Xception uses transfer learning approach for evaluation. As we can see, the most difficult cells to correctly differentiate are the koilocytotic ones in all circumstances. It can be seen that the DenseNet201 is more effective than the other models when it comes to methods that use the classification of cytoplasm and nuclei features. Since the information vanishes before reaching its destination due to the longer path between the input and output layers, DenseNet was designed specifically to improve the accuracy of high-level neural networks that suffer from vanishing gradient problem. Furthermore, we can see that cytoplasm features have a stronger discriminative ability than nucleus features. Figure 5 shows a DenseNet201 feature map with unnecessary features removed and just the most relevant ones included. Table III shows the accuracy results for the four models as before fine-tuning the models and after fine-tuning the models. Densenet201 achieved the highest accuracy, precision, recall and F1-score before and after fine-tuning the model. Equations 2 to 5 include the formula for calculating performance. True positives, true negatives, false positives, and false negatives are all abbreviated as TP, TN, FP, and FN.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

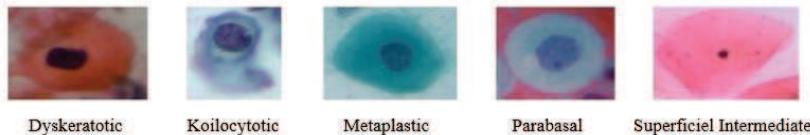


Fig. 2. Sample images of SIPaKMeD dataset

TABLE III
THE SIPaKMeD IMAGE CLASSIFICATION RESULT BEFORE AND AFTER FINE TUNING

Model	Before Fine-Tuning		After Fine-Tuning		Precision (%)	Recall (%)	F1-Score (%)
	Train Accuracy (%)	Test Accuracy (%)	Train Accuracy (%)	Test Accuracy (%)			
InceptionResnetV2	87.66	88.57	96.91	95.58	94.8	95	94
VGG19	70.62	78.21	96.95	94.91	95	94.6	95
Xception	87.96	87.30	92.80	93.31	92.4	93	92
DenseNet201	88.72	90.37	96.63	95.79	95.7	96	96

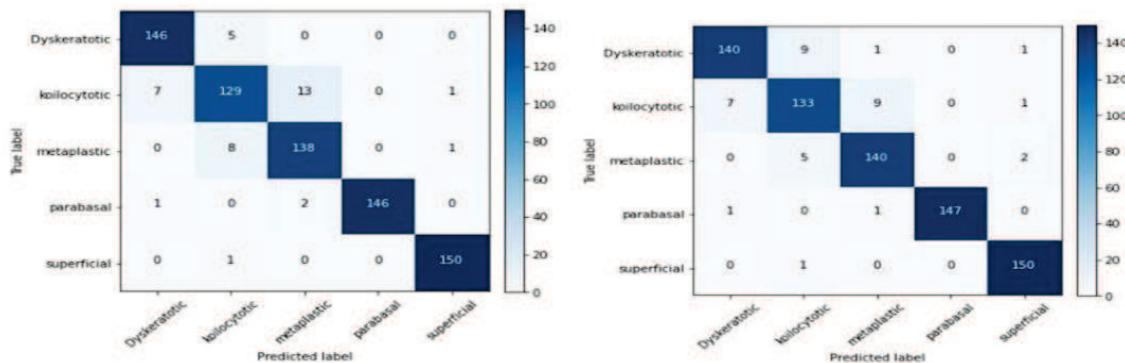


Fig. 3. Confusion matrix of InceptionResNetV2 and VGG19

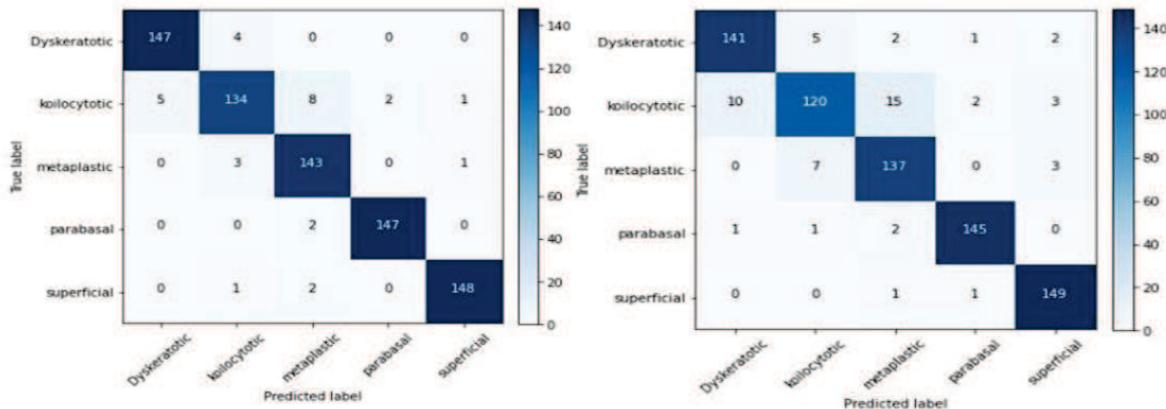


Fig. 4. Confusion matrix of DenseNet201 and Xception

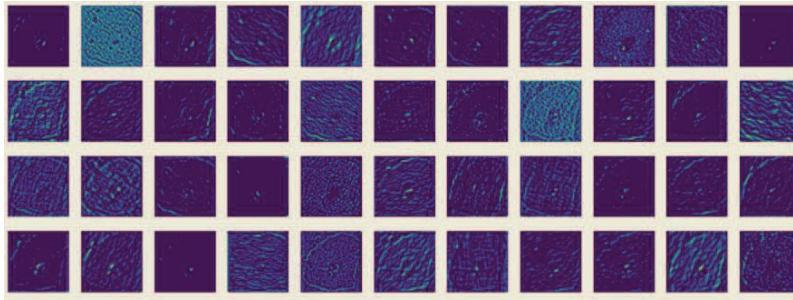


Fig. 5. Feature Map of DenseNet201 Activation layer

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Classifier performance was further studied using information on the confusion matrix of the InceptionResNetV2, VGG19, DenseNet201, and Xception networks in figure 3, 4, 5 and 6.

V. CONCLUSION

This work presents a study of transfer learning frameworks InceptionResNetV2, VGG19, DenseNet201 and Xception networks pre-trained on ImageNet dataset. The SIPaKMeD dataset is used in this study to evaluate the performance of four pre-trained CNN models. Compared to three other pre-trained models, the DenseNet201 outperforms them in cervical cancer images classification with the smallest amount of parameters (18 million) with an accuracy of 96.63%.

In future investigations, adding a detection algorithm that accurately detects and analyses only the cervix could provide a more accurate comparison of cervical cancer classification performance. Finding and implementing improved feature selection approaches could reduce the need for human interaction in machine learning, making it easier to compare different model architectures. We anticipate that future research will allow for a more objective evaluation of different model architectures, assisting physicians in selecting appropriate computer-assisted diagnostic tools.

REFERENCES

- [1] M. N. Asiedu et al., "Development of Algorithms for Automated Detection of Cervical Pre-Cancers with a Low-Cost, Point-of-Care, Pocket Colposcope," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 8, pp. 2306–2318, Aug. 2019, doi: 10.1109/TBME.2018.2887208.
- [2] Y. Liu et al., "Automatic segmentation of cervical nuclei based on deep learning and a conditional random field," *IEEE Access*, vol. 6, pp. 53709–53721, 2018, doi: 10.1109/ACCESS.2018.2871153.
- [3] M. E. Plissiti et al., "Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images," *Proc. - Int. Conf. Image Process. ICIP*, pp. 3144– 3148, 2018, doi: 10.1109/ICIP.2018.8451588.
- [4] L. Hu et al., "An Observational Study of Deep Learning and Automated Evaluation of Cervical Images for Cancer Screening," *J. Natl. Cancer Inst.*, vol. 111, no. 9, pp. 923–932, Sep. 2019, doi: 10.1093/jnci/djy225.
- [5] Kurnianingsih et al., "Segmentation and classification of cervical cells using deep learning," *IEEE Access*, vol. 7, pp. 116925–116941, 2019, doi: 10.1109/ACCESS.2019.2936017.
- [6] C. Li et al., "Transfer learning based classification of cervical cancer immunohistochemistry images," in *ACM International Conference Proceeding Series*, Aug. 2019, pp. 102–106, doi: 10.1145/3364836.3364857.
- [7] C. T. Sari and C. Gunduz-Demir, "Unsupervised Feature Extraction via Deep Learning for Histopathological Classification of Colon Tissue Images," *IEEE Trans. Med. Imaging*, vol. 38, no. 5, pp. 1139–1149, 2019, doi: 10.1109/TMI.2018.2879369.
- [8] Y. Song, L. Zhu, J. Qin, B. Lei, B. Sheng, and K. S. Choi, "Segmentation of Overlapping Cytoplasm in Cervical Smear Images via Adaptive Shape Priors Extracted from Contour Fragments," *IEEE Trans. Med. Imaging*, vol. 38, no. 12, pp. 2849–2862, 2019, doi: 10.1109/TMI.2019.2915633.
- [9] J. Deng, Y. Lu, and J. Ke, "An Accurate Neural Network for Cytologic Whole-Slide Image Analysis," Feb. 2020, doi: 10.1145/3373017.3373039.
- [10] J. M. H. Noottout et al., "Deep Learning-Based Regression and Classification for Automatic Landmark Localization in Medical Images," *IEEE Trans. Med. Imaging*, vol. 39, no. 12, pp. 4011–4022, 2020, doi: 10.1109/TMI.2020.3009002.
- [11] K. P. Win, Y. Kitajadure, K. Hamamoto, and T. M. Aung, "Computer-assisted screening for cervical cancer using digital image processing of pap smear images," *Appl. Sci.*, vol. 10, no. 5, 2020, doi: 10.3390/app10051800.
- [12] D. Xue et al., "An Application of Transfer Learning and Ensemble Learning Techniques for Cervical Histopathology Image Classification," *IEEE Access*, vol. 8, pp. 104603–104618, 2020, doi: 10.1109/ACCESS.2020.2999816.
- [13] A. R. Bhatt et al., "Cervical cancer detection in pap smear whole slide images using convNet with transfer learning and progressive resizing," *PeerJ Comput. Sci.*, vol. 7, pp. 1–18, 2021, doi: 10.7717/peerj-cs.348.
- [14] R. Kundu, H. Basak, A. Koilada, S. Chattopadhyay, S. Chakraborty, and N. Das, "Ensemble of CNN classifiers using Sugeno Fuzzy Integral Technique for Cervical Cytology Image Classification," 2021, [Online]. Available: <http://arxiv.org/abs/2108.09460>.
- [15] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, "A fuzzy rank-based ensemble of CNN models for classification of cervical cytology," *Sci. Rep.*, vol. 11, no. 1, pp. 1–18, 2021, doi: 10.1038/s41598-021-93783-8.
- [16] M. M. Rahaman et al., "DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques," *Comput. Biol. Med.*, vol. 136, no. MI, 2021, doi: 10.1016/j.combiomed.2021.104649.
- [17] A. Tripathi et al., "Classification of cervical cancer using Deep Learning Algorithm," *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, no. Iciccs, pp. 1210–1218, 2021, doi: 10.1109/ICICCS51141.2021.9432382.
- [18] S. Yu et al., "Automatic classification of cervical cells using deep learning method," *IEEE Access*, 2021.

Performance Model Evaluation of Seizure Classification Using Statistical Features and Random Forest

S.Ramya¹

¹*Department of Computational Intelligence*

¹*SRM Institute of Science and Technology*

Kattankulathur, India

email: rs8652@srmist.edu.in

M.Uma²

²*Department of Computational Intelligence*

²*SRM Institute of Science and Technology*

Kattankulathur, India

email: umam@srmist.edu.in

Abstract—The recent emergence of AI and ML techniques has made tremendous progress in healthcare. The automation in predicting chronic diseases is more helpful for medical experts to move on to further treatment without wasting their time on diagnosis. One of the most widespread chronic diseases is epilepsy, which occurs when there is a recurrent seizure. This chronic condition affects approximately 70 million people worldwide, of all ages. The prior prediction of epilepsy would change the life of the subject from their poor lifestyle. The main drawback to creating a prediction model for epilepsy is the lack of datasets and their varied features. Researchers are working hard to develop a good prediction model for clinical use. Random Forest Classifier is used for the classification of the generalized epileptic Seizures with the new dataset which is publicly available in Kaggle. In proposed work, Principal component analysis applied for dimensionality reduction in the training dataset. Supervised machine learning algorithm such as Support vector Machine, Random Forest Classifier and J48 classifiers were taken for analysis and its performance were evaluated using T-Test Paired in weka tool. Finally the Random Forest Classifier has produced good accuracy of 98.56%.

Keywords—: Epilepsy, Seizure, Chronic, recurrent, Random Forest Classifier, Support Vector Machine.

I INTRODUCTION

Epileptic Seizure is a temporary event of signs due to extreme neuronal action of the brain [1]. The brain continuously generates some Electrical impulses in an orderly Pattern. A seizure is an abnormality in this electrical functioning of the brain. Epilepsy is diagnosed when there are two or more repeated Seizures. If a seizure arises from a specific area of the

brain, then the initial symptoms reflect the functions of that area. ES attacks are difficult to foresee, and the strength and duration of an attack cannot be predicted. As a result, the event's injuries, and safety issues are a big concern for patients and their families [2]. Generally, the onset of seizure can be categorized as primary generalized seizures and focal onset seizures. Patients with epilepsy's brain activity can be organized into four states: pre-ictal (immediately prior seizure), ictal (during a seizure), postictal (immediately following a seizure), and inter-ictal (in-between seizures) [2]. The EEG is the most preferred tool for recording brain electrical activity. After a comprehensive review of the EEG data, certain differentiating aspects of a seizure activity have been identified. These include i) background disruption, (ii) clearly defined component, (iii) electro cerebral negativity, and (iv) physiological field [3]. In recent years ML has received increased attention in medical applications. A brief review for (i) Acquisition of physiological signals (ii) Data set description (iii) Feature extraction (iv) ML techniques involved in classification for the automation of ES (Epileptic Seizures) along with their challenges and future enhancements are presented. The rest of the paper is organized into Section II Related work has been discussed. In Section III, the proposed methodology is discussed. In Section IV, the experimental results and discussion are discussed. In Section V, the conclusion is discussed.

II. RELATED WORKS

When models are trained with a large number of datasets, they achieve higher accuracy. Datasets for epilepsy prediction are one of the significant drawbacks. Here, we have reviewed the publicly available EEG datasets along with their usage, and performance is analyzed in using the datasets. Das, et.al, [3] proposed an experimental dataset that consists of 150 subjects. The dataset is classified into a seizure and normal. The sampling frequency is 256 HZ and the duration was 30min. Ayodele, K.P.et al., [6] created a new dataset OAUSZI by fusing two publicly available CHB-MIT datasets and the TUSZ dataset. The accuracy for the combination of these two datasets is 74.5%. Savadkoohi, M., et.al. [10] have used the University of Bonn Dataset which is publicly available. Five EEG signals were collected for healthy participants with five epileptic subjects. From 100 channels 60 features were selected to obtain a precise output. Daoud, H. et.al, [17] in their proposed work used the CHB-MIT dataset which is the most widely used dataset in epilepsy prediction is available at [CHB-MIT Scalp EEG Database v1.0.0](#). There are files of 22 subjects in edf format. It is the European data format in which most of the EEG datasets are available. The author extracted 23 features from 256 samples per second. Wang, G et.al. Have proven accuracy of 90.8% with 21 subjects from the Freiburg EEG dataset which is available in <http://epilepsy.uni-freiburg.de/>. Table 1. Shows the available datasets which are widely used in epilepsy prediction along with their performance, Number of subjects, and Number of features extracted.

III. PROPOSED METHODOLOGY

In Proposed methodology, the acquired EEG signals preprocessed using Band pass filter with the values from 0.5Hz to 3Hz in Mat Lab. Normalization, principal component analysis techniques has been used to reduce the dimensionality of the training dataset. To classify seizure or non-seizure, training and test data are split in the ratio of 80: 20.

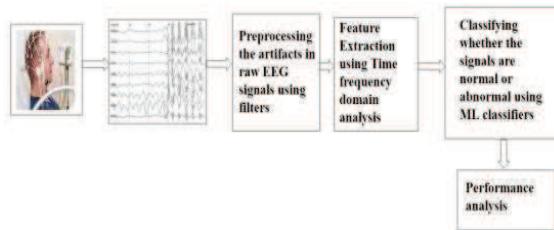


Fig 1. ML approach for Epilepsy Prediction

The RFC, SVM and J48 Classification models has been used to classify the acquired signals and the performance is evaluated. Finally Random Forest Classifier performed well with the greater accuracy of 98.56%. It classifies the subjects with seizure and non-seizure. In, Fig 1 the process involved in the prediction of Epilepsy in the ML approach is depicted.

A. Acquisition of Signal

The signals are obtained through scalp EEG by positioning electrodes on the subject's scalp. Before an experimental setup for acquiring signals the no. of electrodes and the placement of electrodes should be decided based on the targeted potential from which part the signals are acquired. The basic areas of the scalp where the electrodes are placed are listed below.

Electrode Placement:

- 1) Visual Evoked Potentials--Occipital lobe
- 2) Auditory Evoked Potential--Temporal lobe
- 3) Somatosensory Evoked Potentials--Centro parietal
- 4) Control Evoked Potentials--Pre-control Gyrus

After the samples are recorded, a 12-bit analog-digital conversion is performed using an international 10–20 electrode placement scheme based on the data collected. In fig2. The international system of electrode placement is depicted.

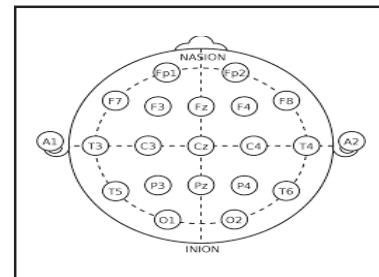


Fig 2.10-20 International System of Electrode Placement

B) Dataset

In this proposed work the publicly available Kaggle dataset updated for a kaggle dataset competition is used for classification. The statistical analysis such as mean, median, mode, standard deviation, min, and max are extracted from the data and the data exploration is done. The monopolar montage of electrode placement is used in this dataset. 22

channels were used for the acquisition of signals. The electrodes are positioned at FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, A1, A2, F2, C2, T1, T2 positions following the international 10-20 system. EEG signals are divided into five spectral components Delta, Theta, Alpha, Beta, and gamma which gives a wide range of frequency components. In this EEG dataset all the spectral components are gathered and analyzed for each channel. Fig 3. Shows the Sample dataset used for predicting seizures.

	FP1_power_delta	FP1_power_theta	FP1_power_alpha	FP1_power_beta	FP1_power_gamma	FP2_power_delta	FP2_power_theta	FP2_power_alpha
15.980332	8.633358	3.092433	0.846559	0.372832	17.581626	9.937238	3.485713	
223.267803	41.573241	13.043525	6.463381	6.168712	298.274432	50.013025	14.411482	
158.673907	31.788608	9.956198	3.736877	2.926979	232.762742	39.332981	10.956081	
29.767326	3.875543	1.553127	0.438266	0.140495	45.206547	5.173542	2.028637	
33.821075	3.635210	1.347280	0.313205	0.080396	55.670049	4.925017	1.783729	

vs x 116 columns

Fig 3. Sample Dataset used for proposed Methodology

C. Preprocessing Methods

The acquired raw EEG signals are pre- processed 1) to decrease the noise 2) Feature Extraction for reducing large amounts of data and 3) for decreasing false positives by various filtering and denoising techniques [7]. Different filters are used in noise distortion including Basic filters, Spatial filters, Wavelet- based techniques, and Adaptive filters. The four Basic filters are used by many researchers but the specific signals cannot be acquired from the overlapped signals by using these basic filters [4]. For data preprocessing, we have used Anaconda Navigator. The missing values from the dataset are identified using two categorical variables as Epileptic and non-Epileptic. The predicted categorical values are 0 and 1. The dataset contains EEG brain signals of 2216 occurrences from 22 features. We classify 1196 events as non-epileptic and 1020 as epileptic. Our main goal is to identify Epileptic Seizures from EEG data. Let's define our independent variables (X) and dependent variable (Y) in order to train our model. Fig 4. Shows the no of 0's as epileptic and 1 as non-epileptic occurrences in the EEG dataset. Further the dataset is filtered using Band pass filter with the values from 0.5Hz to 3Hz in Mat Lab.

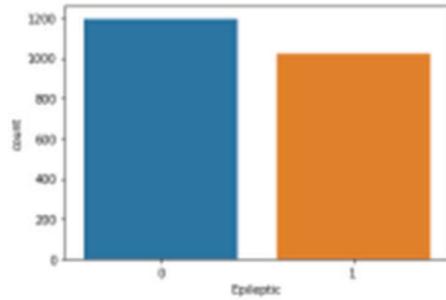


Fig 4. No of Epileptic and Non-Epileptic states

D. Feature Extraction

The exact features are extracted from the preprocessed data to yield a better performance. In our dataset these 22 channels are explored according to their frequency bands as Delta, Theta, Alpha, Beta and gamma so totally we had 110 channels. In our proposed methodology we used statistical analysis to reduce the dimensionality of the features. Time-domain features, Frequency-domain features and Time frequency domain features are the feature extraction methods used to reduce the dimensions in the dataset. In the proposed work features are extracted based on statistical characteristics such as mean, standard deviation, variance, skewness and Kurtosis. The loss of symmetry is measured by skewness. If a distribution or collection of data looks the same to the left and right of the central axis, it is said to be symmetric. The Equations used in statistical moments are as follows:

$$\text{Mean } M(x) = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

$$\text{Variance } V(x) = \frac{1}{n-1} \sum_{i=1}^n (X_i - M(x))^2 \quad (2)$$

$$\text{Skewness } S(x) = N \left[\left(\frac{X-M(x)}{\sqrt{V(x)}} \right) \right]^3 \quad (3)$$

$$\text{Kurtosis } K(x) = N \left[\left(\frac{X-M(x)}{\sqrt{V(x)}} \right) \right]^4 \quad (4)$$

X_i is the i^{th} event in The Data Set n is the number of occurrences in The Data Set. $N(x)$ is the Average of n data points. Using Weka tool further extraction is done using Principal component analysis. In PCA the dimensionality reduction is accomplished by choosing enough eigenvectors. From 550 attributes the dimensions are reduced to 106 attributes. In our proposed methodology, three different classifiers such as Random Forest classifier, Support Vector Machine and J48 were used for classification with a new Kaggle Competition dataset and performance were evaluated based on the accuracy, sensitivity and specificity. Al Mustafa, K.M et.al. in his paper made a performance evaluation on classifiers like Random Forest classifier (RFC), K- Nearest Neighbor (K-NN), Naïve Bayes, Logistic Regression, Decision Tree (DT), Random Tree, J48, Stochastic Gradient Descent (S.G.D.), and showed Random Forest classifier has outperformed of all classifiers with an accuracy of 98.56% in predicting ES [14]. RF classifier is a powerful supervised learning algorithm. It quickly identifies significant information from large datasets aggregated from various decision trees [15]. [16] discussed the various Machine learning algorithms

used to classify ictal & inter-ictal signals and concluded that Genetic algorithms, Bayesian net, fuzzy clustering are not popular classifiers in EEG signal processing. RFC, ANN & KNN are the classifiers that show great accuracy. Even though SVM has some challenges it is the most commonly applied classifier. In our proposed work comparing all three classifiers Random forest classifier outperformed with the greater accuracy of 98.56% with specificity of 97.5%. [13] Analyze the classification model in ML approach with code and performance which makes it easy for the researchers to select the model for predicting epilepsy.

IV Experimental Results and Discussions

In our proposed methodology, recently updated Kaggle datasets were used and the performance of three classifiers was analyzed in order to choose the best ML classifier for predicting ES. In most of the models, SVM classifiers were used and they outperformed for the existing datasets but when considering large datasets, the performance level is minimum than both the RFC and the J48. The recently updated dataset in Kaggle consists of 22 channels the spectral bands of electrodes are analyzed and the statistical data is explored. The Min, Max, Mean, Standard deviation, skewness and kurtosis are calculated for each electrode. Figure 5 shows the data exploration of a single electrode.

	FPI1_power_delta	FPI1_power_theta	FPI1_power_alpha	FPI1_power_beta	FPI1_power_gamma	FPI2_power_delta	FPI2_power_theta
count	2216.000000	2216.000000	2216.000000	2216.000000	2216.000000	2216.000000	2216.000000
mean	5784.616464	968.765238	746.234771	704.487794	1801.437199	2085.575266	992.570476
std	27119.740834	1219.665703	12016.849373	10658.946272	14544.502339	20870.847916	12518.098105
min	0.020945	0.012101	0.009409	0.007828	0.006412	0.026174	0.018836
25%	25.524041	3.983245	1.663825	0.789355	0.381195	27.138890	4.254590
50%	68.865791	8.807711	3.874806	1.795211	0.972968	73.485311	9.688978
75%	271.174096	30.093498	12.052042	6.926583	3.777847	288.354575	29.502963
max	859893.816200	350681.120900	403916.724100	291058.725000	285873.572400	604373.606700	546870.948100

Fig 5. Statistical Data Analysis for the Epilepsy Prediction

In the proposed work J48 and SVM classifier was also analyzed for ES prediction with the same dataset. Weka tool was used for prediction. The absolute mean for each electrode is calculated and the significant electrode is selected based on the threshold value. The accuracy was 89.07% for J48 classifier and for SVM was 88.13%. Since the dataset was too large the accuracy for SVM model is low when compared to J48 and RFC.

a) Performance Evaluation

In this section we discuss the performance of the Random forest classifier when using cross-

validation method for training and testing 80% and 20% respectively. The taken to test our proposed RFC model was 0.11 sec. The performance metrics used to evaluate the model are True positive Rate, False positive Rate, Precision, Recall, F-measure, MCC, ROC Area, and PR Area. The main challenge we indulged in classifying EEG signals was the features in the dataset was too large and feature extraction should be done more effectively to choose the most significant feature. In the total 2216 occurrences 2190 were correctly classified. The root mean squared error was 0.14. A confusion matrix is an error matrix used to visualize a classifier's performance. Table I shows the confusion Matrix for the Random Forest classifier with True positive, True negative , False Positive and False negative values with which the sensitivity and specificity is calculated.

Table I Confusion Matrix for RFC Model

Confusion Matrix	Predicted outcome		Total
	Epilepsy (1)	Non-Epilepsy (0)	
Actual Epilepsy State			
	1	0	
1	2190	15	2205
0	6	5	11
Total	2196	20	2216

The Receiver Operating Characteristic Curve is used to predict the probability of a binary classifier. Figure 6 shows the False Positive Rate on the X-axis and the True Positive Rate on the Y-axis for the proposed RFC model. In Table III a comparison of prediction results along with the dataset and algorithm used is shown. Even though the existing models produce higher results the same publicly available dataset is widely used in existing Methods. In the proposed methodology novelty of a new dataset is taken and the model is trained with the Random Forest Classifier, SVM and J48 and checked for accuracy. In Table II the performance of the RFC model is shown.

Table II Performance metric for RFC

Precision	F-measure	ROC Area	Recall
0.994	0.984	1.00	0.975
0.979	0.987	1.00	0.995
0.986	0.986	1.00	0.986

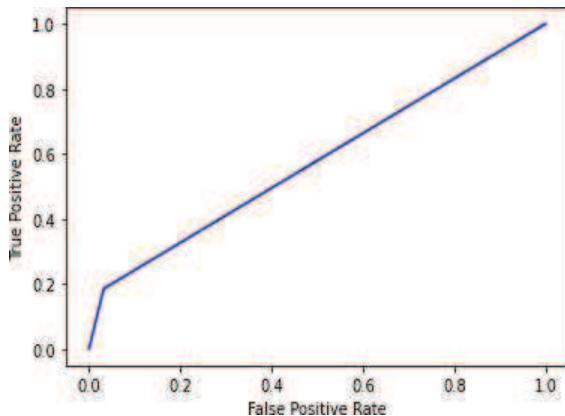


Fig 6. ROC Curve result with true positive rate and false positive rate.

Table III Comparison of Classification Model with different dataset

Ref	Dataset	Algorithm	Accuracy
[18]	University of Bonn	RFC	97.82%
[14]	Freiburg dataset	RFC	90.08%
Proposed Method	Kaggle Competition Dataset	RFC	98.56%

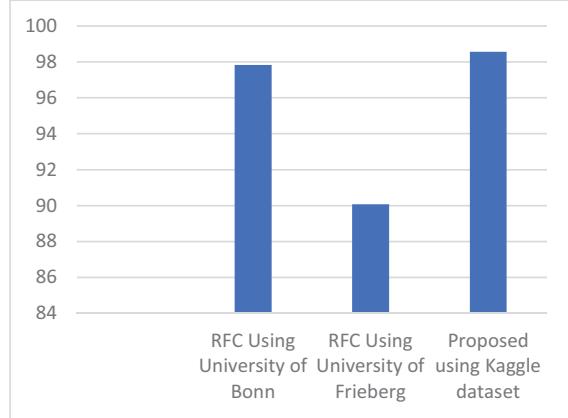


Fig 7. Comparison of prediction models with Existing Datasets and proposed dataset

b) Model Evaluation

In this section, the supervised algorithms such as J48, SVM and RFC are evaluated using t-test paired in weka tool. With the following performance metrics their accuracy is analyzed. The suitable model for classifying Epileptic seizures is selected based upon the evaluation of the model.²²¹⁶ Instances were taken along with 106 attributes for classification and 10 iterations were taken for all algorithms with the same dataset. Table IV shows the comparison of the algorithm with the evaluation metrics

Table IV Model Evaluation Metrics for SVM, J48 and RFC

Evaluation Metrics	J48	SVM	RFC
Correctly Classified Instances	89.07%	88.13%	98.56%
Incorrectly Classified Instances	10.92%	11.86%	1.44%
Kappa Statistic	0.78	0.75	0.97
Mean absolute Error	0.12	0.11	0.17

V Conclusion

Epilepsy begins with a widespread of electrical discharges that involve both sides of the brain at once. EEG the efficient modality that records brain activity is tough to analyze by experts and consumes more time. ML plays an important role in the automation of this analysis. The suggested models have shown higher accuracy with the most widely used dataset. In this study, a new dataset is taken and a comparative analysis of three ML models for predicting ES is done. In the proposed methodology we found that RFC have produced higher accuracy of 98.56% than both the models but since the EEG data is too large, the time taken for training the datasets with ML models is very high. Even though ML techniques are robust in predicting, the efficiency can be improved by DL techniques which have the advantage of automatic feature extraction and predict well for large datasets.

REFERENCES

- [1] Fisher RS, Boas WVE, Blume W, Elger C, Genton P, Lee P, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*. (2005) 46:470–2. DOI: 10.1111/j.0013-9580.2005.66104.x
- [2] Rasheed, K., Qayyum, A., Qadir, J., Sivathamboo, S., Kwan, P., Kuhlmann, L., O'Brien, T., and Razi, A., 2020. Machine learning for predicting epileptic seizures using EEG signals: A review. *IEEE Reviews in Biomedical Engineering*, 14, pp.139-155
- [3] Das, K., Daschakladar, D., Roy, P.P., Chatterjee, A. and Saha, S.P., 2020. Epileptic seizure prediction by the detection of seizure waveform from the pre-ictal phase of EEG signal. *Biomedical Signal Processing and Control*, 57, p.101720.
- [4] Kim, T., Nguyen, P., Pham, N., Bui, N., Truong, H., Ha, S. and Vu, T., 2020. Epileptic seizure detection and experimental treatment: a review. *Frontiers in Neurology*, 11, p.701.
- [5] Zhou, D. and Li, X., 2020. Epilepsy EEG signal classification algorithm based on improved RBF. *Frontiers in Neuroscience*, 14, p.606.
- [6] Ayodele, K.P., Ikezogwo, W.O., Komolafe, M.A. and Ogunbona, P., 2020. Supervised domain generalization for integration of disparate scalp EEG datasets for automatic epileptic seizure detection. *Computers in Biology and Medicine*, 120, p.103757.
- [7] Usman, S.M., Khalid, S., Akhtar, R., Bortolotto, Z., Bashir, Z., and Qiu, H., 2019. Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: Review of available methodologies. *Seizure*, 71, pp.258-269.
- [8] Robbins, K.A., Touyan, J., Mullen, T., Kothe, C., and Bigdely-Shamlo, N., 2020. How sensitive are EEG results to preprocessing methods: a benchmarking study. *IEEE transactions on neural systems and rehabilitation engineering*, 28(5), pp.1081-1090.
- [9] Boonyakanont, P., Lek-Uthai, A., Chomtho, K. and Songsiri, J., 2020. A review of feature extraction and performance evaluation in epileptic seizure detection using EEG. *Biomedical Signal Processing and Control*, 57, p.101702.
- [10] Savadkoohi, M., Oladunni, T. and Thompson, L., 2020. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. *Biocybernetics and Biomedical Engineering*, 40(3), pp.1328-1341.
- [11] Diykh M, Li Y, Wen P. EEG sleep stages classification based on time-domain features and structural graph similarity. *IEEE Trans Neural Syst Rehabil Eng.* (2016) 24:1159–68.doi: 10.1109/TNSRE.2016.2552539.
- [13] Yu, J., 2019. Epileptic Seizure Classification ML Algorithms.
- [14] Almustafa, K.M., 2020. Classification of epileptic seizure dataset using different machine learning algorithms. *Informatics in Medicine Unlocked*, 21, p.100444.
- [15] <https://towardsai.net/p/machine-learning/why-choos%20e-random-forest-and-not-decision-trees>.
- [16] Saminu, S., Xu, G., Shuai, Z., Abd El Kader, I., Jabire, A.H., Ahmed, Y.K., Karaye, I.A. and Ahmad, I.S., 2021. A Recent Investigation on Detection and Classification of Epileptic Seizure Techniques Using EEG Signal. *Brain Sciences*, 11(5), p.668.
- [17] Daoud, H. and Bayoumi, M., 2019. Deep learning approach for epileptic focus localization. *IEEE transactions on biomedical circuits and systems*, 14(2), pp.209-220.
- [18] Li, M., Chen, W. and Zhang, T., 2017. Application of MODWT and log-normal distribution model for automatic epilepsy identification. *Biocybernetics and Biomedical Engineering*, 37(4), pp.679-689.

An Efficient Key Agreement And Anonymous Mutual Authentication Protocols For Secure Communication In VANETs

Swathi Kona
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
konaswathi958@gmail.com

Pawan Kumar Pinninti
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
pawanram963@gmail.com

Sai Venkata Krishna Reddy Morthala
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
morthalasvkreddy0504@gmail.com

Hemanth Kumar Mavuru
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
hemanthmavuru@gmail.com

Rakesh Konathala
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
rakeshkconathala1234@gmail.com

Azees Maria
dept. of ECE
GMR Institute of Technology
(affiliated to JNTUK, KAKINADA)
Rajam, India
azeesmm@gmail.com

Abstract— In recent years, the technology of Vehicular Ad Hoc Networks (VANETs) has turned into a significant field of research. A VANET is a form of network that is built on the notion of constructing a network of vehicles for a given scenario, and it is ad-hoc in nature. Data can be transmitted via vehicle-to-vehicle (V2V) and vehicle-to-roadside (V2R) connections. A VANET is a type of network that is essential for a variety of traffic safety applications and accident prevention. A VANET is a wireless network that connects a group of moving or static vehicles. In vehicle contexts, VANETs play a critical role in delivering safety and comfort to drivers. These include intelligent traffic control, real-time data, and activity management. The data is transmitted through a wireless network. This can lead to security and privacy problems. We proposed an authentication protocol for secure communication in VANETs. Vehicle authentication has been identified as one of the security techniques in vehicular networks for preventing unauthorized access by attackers. Furthermore, the majority of solutions are focused on authentication between V2V communications only. However, V2R authentication is also very much essential for securing VANETs. According to an evaluation of its performance, our suggested system is more efficient in minimizing computation and storage overheads.

Keywords—anonymous credentials, authentication, security, VANETs.

I. INTRODUCTION

The exponential development in the data produced by automobiles has resulted from the tremendous increase in the volume of vehicles on the road, as well as the widespread demand for services by users. Although these significant distinctions between the various domains have emerged, the evolutionary approach for VANET has deviated drastically from the research required for Mobile Ad-hoc Network (MANET). While VANET was once considered a subset of MANET, it is now regarded as a field of study, with a significant increase in the development of VANET-specific applications. As a result, in vehicular networks, cloud technology has the potential to obtain, interpret, and distribute actual data. To minimize data communication delays and avoid performance degradation, efforts have been made to create separate cloud storage rather than cooperate with several Cloud Service Providers (CSPs). As a result, multi-cloud environments have emerged in the automotive system. In several cloud environments, a two-layer vehicular network

layout is used. Vehicles and access points make up the lowest tier. A Trusted Authority (TA), as well as CSP with various support functionality, make up the higher layer. The wireless transmission allows the entities to communicate with one another. VANETs are a sort of ad-hoc network in which vehicles act as nodes for information delivery. Vehicles in the VANET ecosystem include an On-Board Unit (OBU) that enables users to communicate across vehicular nodes using standards including 802.11p, 3G/4G, and others. [1]. The most common modes of communication in VANETs are V2V and V2R. For each of these, communications are employed with the Dedicated Short-Range Communication (DSRC) standard.

The three important components of a basic VANET architecture are TA, Road Side Units (RSUs), and vehicles. The main significant network operator, wireless carrier, and secure credential provider in the VANET system is TA. As a consequence, TA handles important system tasks such as control parameter distribution, user login, vehicular cell layout, user administration, and necessarily connected vehicle verification. It's worth noting that the TA side collects and analyses massive amounts of vehicle data from all legitimate VANET organizations. Massive compute and storage capacities are required. Modern communication and processing techniques, such as the promised 5G technologies and cloud-based services, have been applied to IoT networks, such as VANETs, where sufficient computation and storage abilities can be ensured. A typical VANET system includes OBUs installed in vehicles, RSUs deployed across roadways, and TA [2]. Every 100–300 milliseconds, each vehicle transmits congested statistics such as geolocation, road accident data, and so on to nearby vehicles and RSUs, according to the DSRC's knowledge and prior. By broadcasting information that reflects on-road incidents, RSUs may be able to assist in traffic control. On-road drivers benefit from such messages since they disseminate information about the driving environment. When there is a shortage of adequate infrastructure, an OBU-based authentication method allows vehicles to verify each other for V2V communication. Each vehicle can generate a pair of private key pairs that must be confirmed by other vehicles using a trusted authority-provided private key. The notion of grouping was developed to decrease communication costs. Vehicles that are nearby form a team and elect a team leader, who is in charge of producing a team-shared and secret key pair for communication with the team

leader. The group key could be obtained by a vehicle verifying itself with the group leader. On the other hand, VANETs have a variety of confidentiality problems depending upon the nature of the free and unprotected wireless terminal network links as well as the quick movement of the vehicles. Furthermore, a user's privacy and confidential information, such as true identity-related records, must be securely safeguarded against espionage and violent threats. As a result, for VANETs to be secure, a robust technique for certifying only qualifying vehicles is required. Moreover, because of the asset OBUs, the rapid motion of the vehicles, and the real-time creation and verification of messages, the VANET system has challenges in terms of not only safety and privacy but also transmission and processing weights.

In practical VANET applications, crucial data sharing of V2V and V2R links are handled in an unsecured network domain. Malicious users may intercept communicated vehicular information, exposing the system to a variety of security dangers and privacy problems. Adversaries may gain access to sensitive keying information and user secrets without permission. In this approach, the whole VANET system could be threatened [3]. In this context, dependable methods for maintaining security and protecting confidentiality in VANETs are essential. The authentication process amongst vehicles and RSUs is carried out, along with the access control procedure to validated vehicles, in a variety of schemes using various safe methodologies and algorithms. The particular RSUs are made to distribute the associate's public keys to vehicles in their immediate surroundings. It's all happening through the public channel. However, because of the allotted group key's inherent high mobility, it may be updated at any time, causing severe inference to normal V2R data interchange. To be specific, if any units are deactivated, the incumbent group key must be changed instantly. Simultaneously, the recently included vehicles must be accompanied by the group key. The current VANET security techniques are primarily focused on authentication and efficient key management, with little further research into group membership monitoring approaches. Furthermore, proper and accurate automobile statistics on the vehicle side are critical for customized communications with specific organizations in V2V group interactions. A secure, compatible, and adaptable important management strategy is required due to the enormous number of vehicles registered in various countries and commuting large distances. The fact that legal authorities are involved in automobile registration suggests a degree of centralization. As a result, a Vehicular Public Key Infrastructure (VPKI) is required, in which a TA will provide vehicles with validated public or secret key pairs.

The significant advances in cloud computing technology that have already happened have ushered in new paradigms for massive computing and storage on VANETs. The heterogeneous vehicle data that has already been uploaded could be examined and preserved on a cloud server using sufficient computational resources. In this scenario, adjoining RSUs are known to form a regional vehicular fringe network, whereby typically utilized data could well be kept instantly rather than having to request it from the public cloud every time. These applications have been striving for a distinctive property that a VANET system can provide that hasn't been examined previously, as well as a framework for integrating it efficiently. One of the initial applications for VANET systems was ad hoc, requiring access networks. This system would be able to identify a set of vehicles on its own. The next phase in the evolution of ad-hoc enabled vehicle communities is to use them for vehicle communities formed by rescue teams. When

an emergency service forms a vehicle community, the vehicle community obtains the extra features of being able to drive faster than the maximum speed limit on the road it is on, as well as possessing the right of way in all situations. As a result, they can develop a transmission message scheme that starts from these auto communities and instructs vehicles in front of the group to yield to the unit of emergency service vehicles. It can also be transmitted to side roads that will intersect with the course of the vehicle community to alert such drivers. This device could even make the vehicles move to the side and wait for the emergency services to arrive. The main objectives of the proposed work are listed as follows:

- To develop a reliable key agreement protocol for V2V and V2R communications in VANETs.
- To create a mutual authentication mechanism that is both efficient and secure to avoid unauthorized communication with malicious vehicles in the VANET system.

A. Road map of this paper

This paper is organized as follows. Section II briefly introduces the analysis of literature related to research achievements. Section III represents the System overview. Section IV describes the proposed work in detail. Section V displays the result and discussion. Finally, the conclusion is included in section VI.

II. LITERATURE SURVEY

Y. Yu et al. [4] proposed "Identity-based remote data integrity checking with perfect data privacy preservation for cloud storage" Checking reliability from a distance is a method for a data processing server, such as an internet platform, that demonstrates a validator correctly saving information relating to a data holder. So far, a variety of Remote Intradoc Client (RDIC) procedures have been reported in various studies, but the majority of them suffer from a sophisticated access control problem. In other words, they depend on costly public keys, which may make RDIC implementation problematic in practice. This paper presents a unique personality RDIC method that employs crucial cryptographic techniques to reduce the overall added cost of establishing and maintaining the data encryption authentication framework in RDIC methods. We formalize the data encryption authentication framework RDIC and its authentication process, which incorporates rogue public cloud security and minimal confidentiality from third-party verification. The data encryption authentication framework RDIC protocol that has been proposed does not divulge any personal data. R. Yu et al. [5] proposed "Cooperative resource management in cloud-enabled vehicular networks". The cloud-enabled VANETs are a new revolution for improving the quality of automobile services, and they've gotten a lot of buzz in business and academic circles. In this study, we look at the challenges of bandwidth and processing resource exchange and dissemination in cloud-enabled vehicle networks to support mobile apps. In this situation, CSPs can band together to establish coalitions and share their excess capacity. For CSPs to contribute to existing unused resources, we offer a community strategy focused on two matching concepts. As a consequence, resources may be better allocated and user QoS can be enhanced. When compared to when apps are not cooperative, numerical findings show that our approach will help resource consumption and raise QoS by 75%. Furthermore, the increased service cost of collaboration has a detrimental impact on the creation of coalitions. CSPs' greater desire to collaborate on service applications [6] led them to

propose an "Extensible conditional privacy protection authentication scheme for secure vehicular networks in a multi-cloud environment," which is titled "2020." In a multi-cloud context, few solutions are authentication of vehicular communications with conditional confidentiality as the target. To achieve rapid and effective authentication with CSPs, vehicles Only the TA should be authorized. which relies on cryptographic algorithms instead of complicated bilinear pairing operations. In a multi-cloud scenario, the vehicular network model and certain security goals, assumptions, and design for such a network, Security Objectives V. odelu [7] proposed "A secure biometrics-based multi-server authentication protocol using smart cards". In a multi-cloud scenario, the vehicular network model, as well as certain specific security objectives and constraints for such networks, and Elliptic Curve Cryptography (ECC) with enhanced security features. In a multi-cloud system, few methods focus on the authentication of automobile networks with conditional information privacy, which uses technical security analysis to prove that the method enables secure authentication using the BAN logic. It has a high level of a wide range of capabilities, and other low transmission and computation costs are all provided. As a result, this technology is highly suitable for mobile devices. with good battery life. Y. Liu [8]. "Efficient privacy-preserving dual authentication and key agreement scheme for secure V2V communications in an IoV paradigm." Due to its numerous conditions, this research mainly concentrates on security and privacy. It requires a dual authentication technique. To begin an authenticating session, the OBU produces an unnamed identity and a transient encryption key. Secondly, a TA can verify the authenticity of the real and anonymous characteristics of the vehicle. The vehicle's credibility is then evaluated based on its prior main form of communication, and the V2V session key is established.

M. Ma [9] proposed "an efficient and provably secure authenticated key agreement protocol for fog-based vehicular ad-hoc networks". To develop a unique authenticated key agreement mechanism that does not need bilinear pairing, as well as rigorous comprehensive evidence of safety, and to establish that the proposed scheme satisfies internet VANETs' security needs, S. Bitam [10] proposed "VANET-cloud: A generic cloud computing model for vehicular ad hoc networks." The network access, which is called the cloud computing architecture, intends to represent a large number of computing resources transparently and comprehensively. The current techniques given by intelligent transportation systems in-vehicle networks are mostly focused on increasing the level of safety and ensuring traveler convenience. Cloud computing has a wide range of capabilities for improving ITS. By providing flexible alternatives like different routes, signal timing synchronization, and so on, we may enhance the safety and travel experience. R. I. Meneguette [11] proposed "an efficient resource search and management scheme for vehicular cloud-connected systems." SMART is a mentoring protocol for vehicle mobile cloud searching and tracking that does not require a wayside connection. He also proposed SMART, a mentoring protocol for vehicular cloud platform research and administration that does not require roadside infrastructure. Vehicles must identify and communicate with one another to manage and exchange resources to do this. We compared SMART to two other methods and found that our suggested strategy had a low search time of roughly 0.5 milliseconds for one hop and approximately 0.9 milliseconds for multiple hops. Furthermore, SMART has a high degree of resource availability roughly 87 percent. In the future, we want to modify SMART to operate in metropolitan settings. To

achieve this, vehicles should be self-contained and allowed to communicate with one another to manage and exchange supplies. Z. Wei [12] proposed "Security enhancements for mobile ad hoc networks with trust management using uncertain reasoning" because of the rise in popularity of automobiles, traffic safety, and our daily lives have become intrinsically linked. Many previous studies have looked at the qualities of authenticity, privacy, and traceability in VANETs. Many automobile technologies integrate smart life into their applications. It proposes a fast, seamless, and secure (3-S) messaging method for V2V authentication and communication. In addition to the aforementioned fundamental needs, the 3-S messaging method is intended to decrease computing costs and offer smooth communication when the vehicle travels through different locations. Our scheme may be used in a variety of VANET applications, particularly in safety systems to provide active safety to drivers, thanks to its improved performance and availability. In his paper "Authenticated key agreement scheme for fog-driven IoT healthcare system," X. Jia [13] proposed an "Authenticated key agreement scheme for fog-driven IoT healthcare system." IoT devices, it was argued, are Internet-connected devices that gather data and transfer it to a sorting facility, such as the cloud, for further analysis, filtering, and storage. In plenty of other terms, the data is transferred to a remote data center, which might be situated anywhere in the world, and processed there. Cloud computing offers a lot of drawbacks for applications that require high mobility and latency, especially in hostile environments. These limits can be circumvented to some extent thanks to the fog computing paradigm, which serves as an interface between a faraway data center and an edge device. C.M. Chen [14] proposed "a secure authenticated and key exchange scheme for fog computing". Fog computing is employed in a variety of settings, including smart manufacturing and automotive ad hoc networks. However, inheriting the security concerns of cloud computing is unavoidable as a result of cloud computing. Jia et al. have suggested a haze IoT care system with a secure information exchange technique. Regrettably, we observed that their technique is vulnerable to a transient secret leak attack. [15] who proposed "Dual authentication and key management techniques for secure data transmission in vehicular ad hoc networks". A TA is proposed in this work to deliver a variety of online superior services to customers via VANETs, as well as to conveniently supply a group key to a set of users and keep such a key up to date with defined factors all through users' join and exit operations. A system of key management with two groups is utilized.

M. Azees [16] proposed "Efficient anonymous authentication with a conditional privacy-preserving scheme for vehicular ad hoc networks." The major goal of this study is to offer an effective anonymous authentication mechanism to prevent rogue vehicles from reaching the VANET. Furthermore, the proposed scheme includes a limited level of efficiency and transparency that could be used to track down vehicles or roadside equipment abusing the VANET. P. Sarkar [17] proposed "A simple and generic construction of authenticated encryption with associated data," which suggests that the problem of developing an Authenticated Encryption with Associated Data protocol be addressed (AEAD). A concussion hash function is used with a mechanism in this approach to Authenticated Encryption (AE). Beyond the AE Protocol, the approach is simple and comprehensive, and no more crucial material is needed. D. He [18]. "An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks". A CPPA technique is proposed for vehicular

networks that don't use piecewise linear paring and proves that they can simultaneously offer mutual authentication and privacy protection.

III. SYSTEM OVERVIEW

In this part, the description is about the main physical entities for vehicular ad-hoc networks, such as TA, RSU, and OBU, and then presents some required knowledge for the concept of VANETs. Fig. 1 represents the system model of VANETs.

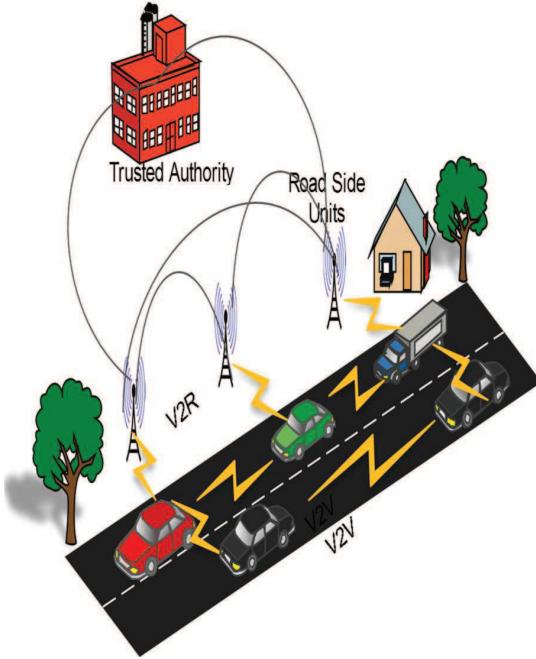


Fig. 1. A system model of VANETs

A. Trusted Authority

The trusted authority is capable of ensuring the VANETs, including registering RSUs, OBUs, and vehicle users. Furthermore, it is responsible for ensuring the security maintenance of VANETs by checking vehicle identification and the unique identity of each user to prevent any vehicle from being involved in accidents or being harmed. There are two primary aspects to trust analysis: direct and indirect trust analysis. Whenever a vehicle tries to access the network with the help of an RSU, the vehicle node is validated using the direct trust evaluation approach. The vehicle node's trust is immediately analyzed based on prior evaluations from the Authority Unit (AU) in addition to deciding if the vehicle component is to control the VANET or not. The TA provides the security events that occur during the interaction between AUs and OBUs are recorded in the database based on past security incident records. The databases of AUs corresponding to similar entities can be shared. It is also used with the recorded occurrences to assess the new vehicle node and determine whether or not it is reliable. When a collection of vehicles forms a wireless connection to transmit data messages with one another, the indirect trust analysis approach can choose to embrace a new vehicle component. The vehicle component's indirect trust evaluation technique is based on the suggested trust from other nodes in VANETs. The indirect trust evaluation allows network nodes to determine whether or not to accept a new access vehicle node.

B. Road Side Unit

There are three elements that are included in RSUs: the transmitter, the CPU, and read-only or write-only memory. The RSUs are placed alongside of the road and are permanently fixed at a specific location from 300m to 1 km away. RSUs have two ways of interfacing: wireless and wired communication. The wireless interface connects to OBUs installed on vehicles, while the cable interface connects between RSUs and the TA. When a prospective accident or emergency braking and traffic condition is identified, alert notifications are sent from the network through RSUs to all vehicles inside its range, improving road traffic and accident prevention, particularly on curved routes, crossroads, and narrow roads. As a result, VANETs' maximum range might span the whole road stretch. In practice, several RSUs are situated in challenging natural settings far from the remote database. As a result, it's conceivable that the RSUs will be deactivated. As a result, the critical vehicle confidential information, as well as the exact vehicle identification, must not be completely exposed to RSUs for privacy reasons.

C. On Board Unit

An OBU is a hardware device mounted on a vehicle without the help of roadside units. The OBU detects its location and analyses the essential data. OBU is made up of vehicle networks that communicate via V2V and V2R. OBUs are equipped with wireless communication modules, including a transmitter. The OBU is made up of the control unit, the communication subsystem, the sensor network, and the transceiver. An RF antenna is connected to a CPU that can also read and write memory to store data in this situation. Data leakage that is undetectable in the vehicle sensor network or during transmission to the monitoring centre can put people's lives at risk. The major responsibility of the OBU is to communicate with other board units and roadside units.

IV. PROPOSED WORK

A. Initialization

TA chooses a random number $q \in z_q^*$ where $z_q^* = (1, 2, 3, 4, \dots, q-1)$. Let g_1 be the element of the group G TA chooses secret keys $m, n \in z_q^*$ and private key $r \in z_q^*$. Based on these keys public keys are calculated as $M = g_1^m$, $N = g_1^{m+n}$ and $R = g_1^r$ respectively. $H(\cdot)$ be the hash function. The public parameters are $(q, g_1, G, M, N, R, H(\cdot))$.

B. Vehicle Registration and Authentication

Initially vehicle should register with the TA through its required credentials. TA chooses a private key(v) for vehicle and calculates the public key for vehicle as $V = g_1^{v-n}$.

The OBU in the vehicle chooses random numbers $a, b, c \in z_q^*$ and calculates V_1 and V_2 respectively where

$$V_1 = V \cdot g_1^{a-v}, \quad V_2 = g_1^{b+c}$$

and calculate the challenger value for vehicle as $ch_v = H(V||M||V_2)$. Moreover, OBU calculate V'_1, V'_2 where

$$V'_1 = g_1^{-a}, \quad V'_2 = g_1^{a+b+c}$$

and sends (ch_v, ϕ, V'_1, V'_2) to RSU

Here, ϕ is message send to RSU.

By receiving these parameters, the RSU calculates V''_1, V''_2 and verifies the challenger values. Then the RSU also calculates ch'_v where $ch'_v = H(V||V''_1||V''_2)$. If $ch'_v = ch_v$ is equal, then the vehicle user is accepted.

proof of correctness:

$$\begin{aligned} V_1'' &= V_1' \cdot V_1 \cdot N = g_1^{-a} \cdot V \cdot g_1^{a-v} \cdot g_1^{m+n} \\ &= g_1^{-a} \cdot g_1^{v-h} \cdot g_1^{a-v} \cdot g_1^{m+n} = g_1^m = M \\ V_2'' &= V_1' \cdot V_2' = g_1^{-a} \cdot g_1^{a+b+c} = g_1^{b+c} = V_2 \end{aligned}$$

C. RSU Registration and Authentication

RSU should register in TA using its required credentials. Then, TA provides RSU with private and public keys as $s \in z_q^*$ and $S = g_1^{s+r}$ respectively.

Furthermore, the RSU chooses some random number $d, e, f \in z_q^*$ and calculates

$$S_1 = g_1^{d-e}, S_2 = g_1^{f-e},$$

and calculate the challenges values for RSU as

$$ch_s = H(S||R||S_1||S_2)$$

Then, the values of S'_1, S'_2, S'_3 in calculated by RSU as

$$S'_1 = g_1^{d-e+f}, S'_2 = g_1^{-f}, S'_3 = g_1^{-d}$$

and it sends $(ch_s, \alpha, S'_1, S'_2, S'_3)$ to vehicle side

Here, α is message send to vehicle

Then, the OBU of the vehicle calculates S''_1 and S''_2 checks for the ch_s (challenger value of RSU). Once $ch'_s = ch_s$, where $ch'_s = H(S'_1||S'_2||S||R)$, the received message (α) is accepted and the RSU is authenticated.

proof of correctness:

$$\begin{aligned} S''_1 &= S'_1 \cdot S'_2 = g_1^{d-e+f} \cdot g_1^{-f} = g_1^{d-e} = S_1 \\ S''_2 &= S'_1 \cdot S'_3 = g_1^{d-e+f} \cdot g_1^{-d} = g_1^{f-e} = S_2 \end{aligned}$$

V. RESULTS AND DISCUSSION

In this part, we assess the proposed scheme's performance with respect to authentication and signature validation computation costs. The following measures were used to compare the existing schemes to the proposed scheme:

A. Proposed Scheme Computation Cost of authentication and Signature Validation

To identify a vehicle and validate the correctness of a communication, it takes the entire duration involved to evaluate signatures and certificates. This scheme's verification delay is longer than that of several other schemes. Pairing-free certificateless authentication scheme [19], PATF [20], EABA [21], LIU [22]. In order to assess the actual calculation time of our proposed scheme, we utilized a 2-GHz system with 8-GB of installed RAM, running Cygwin 1.7.35-15 [23] and GCC version 4.9.2 for our simulations. In addition to PFCA, PATF, EABA, LIU, the proposed scheme can validate the maximum number of signatures and certificates in 300 milliseconds. The proposed scheme takes only 3.9 milliseconds to verify a single certificate and signature, whereas the existing schemes PFCA, PATF, EABA, LIU take 25.5 milliseconds, 13.6 milliseconds, 11.7 milliseconds, 5 milliseconds, respectively. As T_p , and T_h are the most time-consuming functions in the signature validation process. The verification time can be calculated using a certificate that was sent with a message signature. The certificate and signature verification costs for PFCA, PATF, EABA, LIU, and proposed work are summarized in Table I.

TABLE I. CERTIFICATE AND SIGNATURE VERIFICATION COST FOR VARIOUS SCHEMES

Schemes	For n Vehicles	For n Road Side Units
PFCA [19]	$4nT_p + 5nT_h + 6nT_{pm} + 2nT_m + 3nT_{pa}$	$(5n + 3)T_p + (2n + 3)T_h + 2nT_m + 3nT_{pa}$
PATF [20]	$4nT_p + 2nT_h + 3nT_{pm}$	$3nT_{pm} + 2nT_p + 2nT_h + 2nT_m$
EABA [21]	$3nT_{pa}3nT_h + 3nT_{pm}$	$4nT_{pa} + 2nT_h + (2n + 1)T_{pm}$
LIU [22]	$2nT_p + nT_h + nT_{pa} + 2nT_{pm}$	$2nT_p + 2nT_m + 2nT_{pa} + 2nT_{pm}$
PROPOSED WORK	$nT_h + 3nT_{pm}$	$nT_h + 3nT_{pm}$

From Table I. it is shown that the proposed technique has the lowest computational cost among the many current schemes for performing certificate and signature validation since it only needs nT_h , $2nT_{pm}$ to check one certificate and signature.

In Table II, the execution time of Cryptographic operations is represented.

TABLE II. EXECUTION TIME OF OPERATION

Notations	Operations	Execution time (ms)
T_p	Bilinear pairing	1.6
T_{pm}	Point multiplication	2.7
T_h	Hash function	0.6
T_{pa}	Point addition	0.6
T_m	Modular operation	0.1

When n is high, it shown that the current proposed scheme is substantially more efficient than other current schemes and provides the shortest validation time among the schemes under consideration.

From Fig. 2, that our proposed scheme takes just around 390 milliseconds to validate 100 certificates and signatures. Other approaches, require more than 500 milliseconds to validate 100 certificates and signatures.

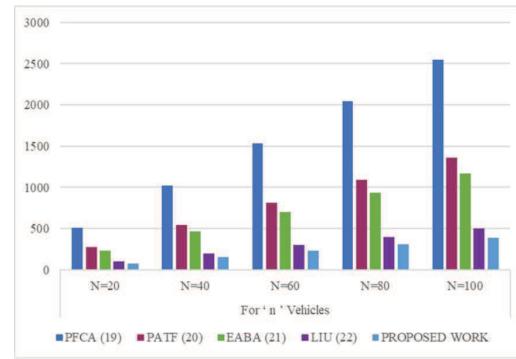


Fig. 2. Computational Cost for n Vehicles

From Fig. 3, it is shown that in our proposed scheme, the RSU can produce 100 certificates and signatures every 450 milliseconds, while the vehicles can validate 200 certificates and signatures every 300 milliseconds, so that the proposed

technique gives the least message loss as the number of vehicles within the communication range expands.

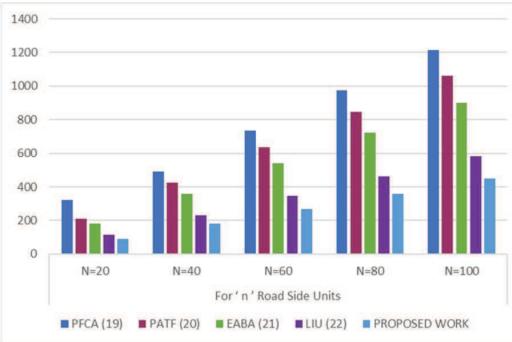


Fig. 3. Computational Cost for n Road Side Units

VI. CONCLUSION

The proposed paper is to provide an outline and explanation of different VANET challenges. In the context of vehicular communication, a variety of research concerns are addressed. This paper proposed efficient authentication and key management protocols to achieve secure communication among various vehicles. Each vehicle must be registered with the TA using the appropriate credentials. TA selects a vehicle's private key and calculates the vehicle's public key. Using these anonymous credentials, the privacy of the vehicle's users is highly preserved. According to the performance evaluation, our system is more accurate than previous schemes in terms of processing and computational cost overhead, making it more suited for deployment in VANET services and applications.

REFERENCES

- [1] Zhang, J., Zhong, H., Cui, J., Xu, Y., & Liu, L. (2021). Smaka: Secure many-to-many authentication and key agreement scheme for Vehicular Networks. *IEEE Transactions on Information Forensics and Security*, 16, 1810–1824.
- [2] Tan, H., & Chung, I. (2020). Secure authentication and key management with blockchain in VANETs. *IEEE Access*, 8, 2482–2498. <https://doi.org/10.1109/access.2019.2962387>
- [3] Alfadhl, S. A., Lu, S., Chen, K., & Sebai, M. (2020). MFSPV: A multi-factor secured and lightweight privacy-preserving authentication scheme for Vanets. *IEEE Access*, 8, 142858–142874. <https://doi.org/10.1109/access.2020.3014038>
- [4] Yu, M. H., Au, G., Ateniese, X., Huang, W., Susilo, Y., Dai, and G. Min, "Identity-Based Remote Data Integrity Checking With Perfect Data Privacy Preserving for Cloud Storage," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 767–778, 2017.
- [5] R. Yu, X. Huang, J. Kang, J. Ding, S. Maharjan, S. Gjessing, and Y. Zhang, "Cooperative Resource Management in Cloud-Enabled Vehicular Networks," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7938–7951, 2015.
- [6] J. Cui, X. Zhang, H. Zhong, J. Zhang, and L. Liu, "Extensible Conditional Privacy Protection Authentication Scheme for Secure Vehicular Networks in a Multi-Cloud Environment," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1654–1667, 2020.
- [7] V. Odelu, A. K. Das, and A. Goswami, "A Secure Biometrics-Based Multi-Server Authentication Protocol Using Smart Cards," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1953–1966, 2015.
- [8] Y. Liu, Y. Wang, and G. Chang, "Efficient Privacy-Preserving Dual Authentication and Key Agreement Scheme for Secure V2V Communications in an IoV Paradigm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2740–2749, 2017.
- [9] M. Ma, D. He, H. Wang, N. Kumar, and K.-K. R. Choo, "An Efficient and Provably Secure Authenticated Key Agreement Protocol for Fog-Based Vehicular Ad-Hoc Networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8065–8075, 2019.
- [10] S. Bitam, A. Mellouk, and S. Zeadally, "VANET-cloud: a generic cloud computing model for vehicular Ad Hoc networks," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 96–102, 2015.
- [11] R. I. Meneguette, A. Boukerche, and R. D. Grande, "SMART: An Efficient Resource Search and Management Scheme for Vehicular Cloud-Connected System," 2016 IEEE Global Communications Conference (GLOBECOM), 2016.
- [12] Z. Wei, "Trust Management for Security Enhancements in Ad hoc Networking Paradigms with Uncertain Reasoning."
- [13] X. Jia, D. He, N. Kumar, and K.-K. R. Choo, "Authenticated key agreement scheme for fog-driven IoT healthcare system," *Wireless Networks*, vol. 25, no. 8, pp. 4737–4750, 2018.
- [14] C.-M. Chen, Y. Huang, K.-H. Wang, S. Kumari, and M.-E. Wu, "A secure authenticated and key exchange scheme for fog computing," *Enterprise Information Systems*, vol. 15, no. 9, pp. 1200–1215, 2020.
- [15] P. Vijayakumar, M. Azees, A. Kannan, and L. J. Deborah, "Dual Authentication and Key Management Techniques for Secure Data Transmission in Vehicular Ad Hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1015–1028, 2016.
- [16] M. Azees, P. Vijayakumar, and L. J. Deboarh, "EAAP: Efficient Anonymous Authentication With Conditional Privacy-Preserving Scheme for Vehicular Ad Hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2467–2476, 2017.
- [17] P. Sarkar, "A Simple and Generic Construction of Authenticated Encryption with Associated Data," *ACM Transactions on Information and System Security*, vol. 13, no. 4, pp. 1–16, 2010.
- [18] D. He, S. Zeadally, B. Xu, and X. Huang, "An Efficient Identity-Based Conditional Privacy-Preserving Authentication Scheme for Vehicular Ad Hoc Networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2681–2691, 2015.
- [19] Gayathri, N. B., Thumur, G., Reddy, P. V., & Ur Rahman, M. Z. (2018). Efficient pairing-free certificateless authentication scheme with batch verification for vehicular ad-hoc networks.
- [20] Jiang, S., Zhu, X., & Wang, L. (2016). An efficient anonymous batch authentication scheme based on HMAC for VANETs. *IEEE Transactions on Intelligent Transportation Systems*, 17(8), 2193–2204.
- [21] Jing, T., Pei, Y., Zhang, B., Hu, C., Huo, Y., Li, H., & Lu, Y. (2018). An efficient anonymous batch authentication scheme based on priority and cooperation for Vanets.
- [22] Liu, Y., Wang, Y., & Chang, G. (2017). Efficient privacy-preserving dual authentication and key agreement scheme for Secure V2V Communications in an IOV paradigm.
- [23] Cygwin: Linux Environment Emulator for Windows. [Online]. Available: <http://www.cygwin.com/>.

A Robust Approach based on CNN-LSTM Network for the identification of diabetic retinopathy from Fundus Images

¹Sajeev Ram Arumugam
Department of AI&DS
Sri Krishna College of Engineering and
Technology
Coimbatore, India
imsajeet@gmail.com

⁴Balakrishna R
Department of CSE
Vels Institute of Science Technology
and Advanced Studies(VISTAS)
Chennai, India
krishna.se@velsuniv.ac.in

²E.Anna Devi
Department of ECE
Sathyabama Institute of Science and
Technology
Chennai, India
annadevimit@gmail.com

³ V.Rajeshram
Department of Computer Science and
Engineering
M.Kumarasamy College of Engineering
Karur, India
rajeshram107@gmail.com

⁵Sankar Ganesh Karuppasamy
Department of CSE
Vel Tech Rangarajan Dr. Sagunthala
R&D Institute of Science and
Technology
Chennai, India
prof.sankarganesh@gmail.com

⁶S.Vinoth Kumar
Department of CSE
Vel Tech Rangarajan Dr. Sagunthala
R&D Institute of Science and
Technology
Chennai, India
profsvinoth@gmail.com

Abstract— Diabetic retinopathy (DR) is a condition that infects the eyes that involve people with diabetes losing their vision. It influences the blood vessels of the eye. Sometimes people complain about eyesight problems, such as difficulties reading or seeing too far away. The retina's blood vessels begin to bleed in the later disease stages. Highly trained experts typically examine coloured fundus images to detect this fatal condition. This condition's manual diagnosis is time-consuming and error-prone. As a result, many computers vision-based algorithms for automatically detecting DR and its various stages from retina images have been presented. We used the Kaggle retina image dataset for this study, which is openly accessible. We introduced a convolutional neural network (CNN), and long short-term memory (LSTM) based deep learning technique for diagnosing diabetic retinopathy. The system attains better performance than the existing systems.

Keywords— *Diabetic retinopathy, CNN, LSTM, deep learning, Kaggle, fundus images.*

I. INTRODUCTION

Diabetic retinopathy is a severe problem that is amongst the primary reasons for visual impairment all around the world, with the incidence of diabetic individuals forecast to rise by 346 million in 2012 to 552 million by 2030 [1]. With early detection and appropriate treatment, the majority of cases of vision loss can be avoided. One-third of diabetic individuals have Diabetic retinopathy without any signs of vision difficulties, allowing the illness to proceed without therapy. As a result, diagnostic protocols and routine checks are crucial for detecting Diabetic Retinopathy at the initial stages. Non-proliferative DR and Proliferative DR are two categories of Diabetic Retinopathy distinguished by the absence of development of new aberrant vasculature. Non-proliferative DR is classified as mild, moderate, or severe [2]. The likelihood of outcome can sort these levels. Mild Non-proliferative DR is the occurrence of microaneurysms (MAs), and it is the primary symptom of Diabetic Retinopathy. As the condition advances, haemorrhages (HEMs) tend to appear in the retina's greater depths. As the disease progresses, MAs, dot-haemorrhages, retinal thickness, hard exudates, intraretinal microvascular abnormalities, venous beading, and cotton-wool spots are all soft exudates that appear. At last, PDR is characterized by

vascular mutation formation. Fig.1c shows diabetic retinopathy stages [3].

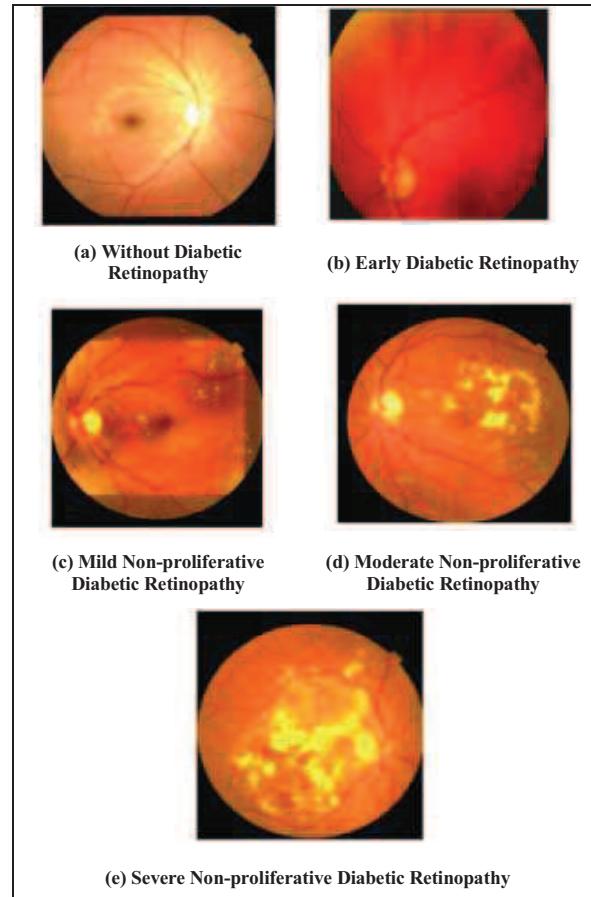


Fig 1. Stages of Diabetic Retinopathy

Computer-aided diagnosis (CAD) technologies can help enhance the Diabetic Retinopathy screening process by lowering the stress on ophthalmologists and offering a second impartial view, minimizing subjectivity in the evaluation. Deep learning, in particular, has recently enabled

these machines to achieve fairly close Diabetic Retinopathy detection ability. DR grading, or the classification of pathology as per severity, is a more complex topic to solve as it involves recognizing and combining several anomalies. Experts sometimes disagree on how to grade complex cases; therefore, the outside perspective of CAD systems could help Diabetic Retinopathy grading become even more efficient. Due to their high complexity and intrinsic black-box behaviour, deep learning systems' decisions are challenging for experts to comprehend and understand [4]. Because of this, we provide a Diabetic Retinopathy categorizing system that provides a medically apparent reason for its conclusions and evaluates the degree of uncertainty in the prognosis. The ophthalmologist understands the system's assessment and how that assessment is received, attempting to make the Diabetic Retinopathy diagnosis less vulnerable.

II. RELATED WORK

As it allows for better use of the enormous amount of data accessible and faster handling of the labelling noise produced by job complexity, deep learning has overtaken field knowledge-based feature engineering as the preferred technique for DR detection. A Diabetic Retinopathy categorizing system backs up its findings with an intelligible description and an assessment, permitting the ophthalmologist to weigh in on the choice [5]. It can understand an image position coupled with an interpretation map and a predictive ambiguity to a unique Gaussian-sampling method based on Multiple Instance Learning frameworks.

Based on active deep learning, automatic recognition of the Diabetic Retinopathy stage is presented (ADL). We employed the CNN model to mine characteristics dynamically. The CNN process requires many labelled data to train, making classification practically hard. A label-efficient CNN architecture is proposed using anticipated gradient length, ADL-CNN [6]. The model had a 92.20 % sensitivity and 95.10 % specificity. As a result, the new ADL-CNN architecture outperforms the older one to detect DR-related lesions. The goal is to use CNN-based transfer learning to implement DR fundus image identification. Investigators are accompanied using the DR1 and MESSIDOR databases, both of which are freely available [7]. We conduct experiments utilizing three different methods to fulfil the objective task: 1) fine-tuning every network layer of numerous pre-trained Convolutional Neural Network models; 2) layer-by-layer fine-tuning of a pre-trained Convolutional Neural Network model; 3) automated extraction from fundus images using pre-trained CNN models

The research goal is to create an automated method for detecting hemorrhages in retinal images. The finite-difference model of the proposed system is obtained by processing an image using variational mode decomposition. A classifier is trained using estimated texture descriptors to differentiate among standard images and hemorrhages images [8]. Histogram equalization, as well as contrast limited adaptive histogram equalization

methods, are used in the proposed approach [9]. A CNN classification is used to make the diagnosis.

A new approach for detecting hard exudates is highly accurate in terms of lesion level. Using the background subtraction methodology, potential candidate exudate lesions are identified. The de-correlation stretch-based method was used to remove the spurious exudate lesion detections. It is tested with the DiaretDB database, which is open to the public [10]. A unique diabetic retinopathy detection using a convolutional network integrates multi-view fundus images for automatic Diabetic Retinopathy detection. The approach fully leverages the retinal lesion features with 120 to 150 degrees field-of-view. More focus will be directed to the prominent view due to the introduction of attention mechanisms, and performance will increase [11].

The entropy image of fundus image is proven to raise the recognition accuracy for Diabetic Retinopathy using CNN. Unsharp masking is also utilized for computing the images. A bichannel CNN combining the entropy images properties of the grey level as well as the green component is presented to create a more accurate performance [12]. A Multiple Instance Learning (MIL) framework was developed, which takes advantage of the inherent information contained in image-level annotations. The main contribution is strengthening the occurrence encryption and classification techniques. A loss function that assures suitable occurrence and multi representations improve system outcomes comprehensibility. [13].

III. MATERIALS AND METHODS

A. Dataset

The publicly available "Kaggle Diabetic Retinopathy" dataset contains 35,126 colour fundus images with sizes ranging from 433,289 pixels to 5184,3456 pixels, which were collected using different fundus cameras in many eye facilities across California and the United States [14]. Colour fundus images with good image quality are chosen. The International Disease Severity Scale was used to rate the images of the retina from the Kaggle dataset: no evidence of retinopathy (stage 0), Mild Nonproliferative DR (stage 1), Moderate Nonproliferative DR (stage 2), Severe Nonproliferative DR (stage 3), and Proliferative DR (stage 4).

B. Convolution Neural Network

A convolutional neural network (CNN), often called as ConvNet, is a deep learning method that uses images as input and uses learnable weights to distinguish one image from another. Compared to other traditional techniques, CNN requires significantly less preprocessing [15]. The convolutional neural network (CNN) is an architecture inspired physically and conceptually by the hierarchical arrangement of the human visual brain. The training is strong enough to provide correct predictions despite the exact position of deformed, transformed, and constant object. Data augmentation is used to acquire input invariances and increase the model's localization capacity. Investigators adapt and modify several CNN designs to improve their application in medical imaging. A ConvNet can extract spatial and temporal image dependencies by using

appropriate filters. The layers that make up the CNN are depicted in Fig.2.



Fig.2 Convolution Neural Network Architecture

It has three layers: an input layer, a hidden layer, and an output layer. The image data is stored in the input layer. The convolutional layer is the primary step to extract a feature from an input image, and it may also include the activation layers. Between two convolution layers, the pooling layer is always present and is used to minimize the spatial volume of the data [16]. This layer contains only hyperparameters. Fully linked layers connect neurons in one layer to neurons in other layers and hold weights and biases. It is the one who is in charge of using training to classify images. The SoftMax layer for multi-classification or the logistic layer for binary classifiers is the final layer. The classification label is stored in the output layer. The convolution layer works as follows:

$$C(a, b) = (A * K)(a, b) = \sum \sum A(a+i, b+j)K(i,j) \quad \dots(1)$$

Where A is the input matrix, K indicates a size filter (a,b), and C represents feature map output. ($A * K$) represents the convolution layer output—the pooling layer downsamples a specific input dimension to limit the number of parameters. The standard technique is max pooling; it generates a higher value in an input region. The Fully Connected layer makes decisions by considering features from the convolutional and pooling layers.

C. Long Short Term Memory

To deal with the issue of fading and inflating gradients, LSTM introduces memory chunks. Long-term dependencies can be handled with LSTMs. LSTMs can recall and relate information from the past to the present [17]. In the LSTM, a memory block is a multi-memory cell functional unit consisting of more memory cells. A couple of multiplicative gates are used as input and output gates. A collection of adaptable multiplicative gates regulates the entire operation of a memory block. The input gate executes an accept or reject process for a cell activation input flow to a memory cell [18]. The output gate controls whether a memory cell's output state is allowed or rejected by other nodes. Fig.3 shows LSTM architecture.

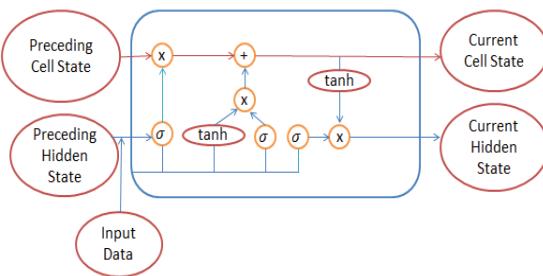


Fig. 3 LSTM Architecture

The LSTM Architecture works as follows.

- The forget gate is provided with the initial cell state from the preceding moment and the current time's input data, and the forget gate's final result is obtained. The forget gate selects what bits of long-term memory are to be remembered based on the previous hidden state and the current data point.
- The goal is to determine which new data should be stored in the network's long-term memory, assuming the past hidden state and incoming input data. It is sent into the new memory network and the input gate.
- This filter should be applied to the cell that has recently been updated. It assures that the required data is produced. When utilizing the filter, we run the cell state through a tanh to force the values within the [-1,1] range.

D. CNN-LSTM Network

A combination technique was established in this work to diagnose diabetic retinopathy automatically. This architecture structure was created by integrating CNN and LSTM networks, with the CNN extracting complicated features from images and the LSTM acting as a classifier, as shown in Fig. 4.

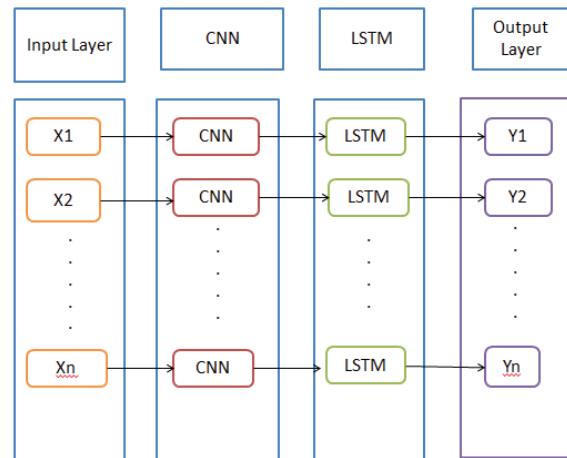


Fig. 4 Architecture of CNN-LSTM network

Our model's architecture consists of an input layer, four Convolution Neural Network layers enfolded in time-distributed layers, an LSTM layer, a dense layer, a dropout layer, and an output layer. Before building the model, we double-checked that all data had been translated into a readable form. The padded sequences were then transformed into an $m \times n$ matrix using one-hot encoding as the following vectorization step. The network was ready for feature extraction after information was padded and vectorized. The relu activation function is used to extract features in this stage. 128 filters were employed in these convolutional layers. We used a flatten layer to combine all retrieved features before delivering them to the LSTM layer. The softmax activation function was utilized to define binary classification outputs. The testing dataset keeps track of how far the training phase has progressed so that the model's

training can be completed sooner if the convergence differs. We used categorical cross-entropy to have two or even more one-hot encoded label classes. For each momentum-like parameter, Adam generates adaptive learning values. It improves multi-class classification models using a softmax activation function.

IV. RESULTS AND DISCUSSIONS

The first and most critical stage of our research is data preprocessing. Indeed, if the data is not prepared correctly, the network will be trained wrongly, making it impossible to generate reliable results. As a result, after the encoding, padding, and vectorization processes, we double-checked all outputs. Another challenge in the model design is combining the multiple neural networks. Incorrect network design can enhance bias and lead to unexpected outcomes. As a result, we ensured that all levels were appropriately cascaded.

We developed a hybrid technique for the identification of DR utilizing CNN as well as LSTM. We started by transforming the original sequences to vectors using "one-hot" encoding to prepare the dataset. After that, all of the data was padded and vectorized. The relu activation function was required to obtain features using three convolution layers. We used a flatten layer to combine every retrieved feature before delivering them to the LSTM layer. The softmax activation function was employed to provide outputs for binary classification.

There are a variety of performance metrics that can be used to assess the classification performance of Deep Learning algorithms. Accuracy, sensitivity and specificity are all common Deep Learning measurements. The percent of accurate predictions for the sample data is known as accuracy. The percent of aberrant images identified as abnormal is sensitivity, whereas the percent of standard images identified as normal is specificity. The percent of properly categorized images is known as accuracy. The formulae for each measurement are listed below.

$$\text{Accuracy} = \frac{\text{Accurate Positivity} + \text{Accurate Negativity}}{\text{Accurate Positivity}} \quad \dots(2)$$

$$\text{Sensitivity} = \frac{\text{Accurate Positivity}}{\text{Accurate Positivity} + \text{Inaccurate Negativity}} \quad \dots(3)$$

$$\text{Specificity} = \frac{\text{Accurate Negativity}}{\text{Accurate Negativity} + \text{Inaccurate Positivity}} \quad \dots(4)$$

The proposed system attains an accuracy of 96.3%, sensitivity of 95.8% and specificity of 95.3%, which is better than the existing system. AUC is a composite performance statistic that considers all available classification levels. AUC is what percentage of the system rates a randomized positive example better than a randomized negative one. Fig. 5 shows the AUC curve of the proposed model.

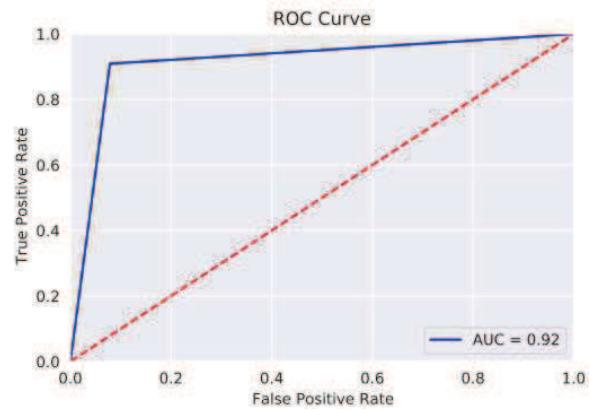


Fig 5 AUC Curve

V. CONCLUSION

In this research, we develop a method for detecting diabetic retinopathy by combining CNN and LSTM Network. During the preprocessing step, one-hot encoding transforms every base into an identical size matrix. Three convolution layers were used, each enclosed in a time-distribution layer. We used a flatten layer to concatenate all retrieved features before delivering them to the LSTM layer. We created an LSTM layer after a dropout layer. The softmax activation function is employed to determine the outputs for binary classification. Our findings revealed that the presented technique is widely trained on the training dataset and outperformed the earlier versions of the testing data. The outcomes showed that the combined Convolution Neural Network and Long Short Term Memory networks had improved diabetic retinopathy classification performance than the existing systems and accuracy of 96.3%, sensitivity of 95.8% and specificity of 95.3%.

REFERENCES

- [1] T. Scully, "Outlook Diabetes: Diabetes in Numbers," *Nature*, vol. 485, no. 7398, p. S2, 2012.
- [2] F. Ghanchi *et al.*, "The Royal College of Ophthalmologists' clinical guidelines for diabetic retinopathy: A summary," *Eye*, vol. 27, no. 2, pp. 285–287, 2013, doi: 10.1038/eye.2012.287.
- [3] J. P. Kandhasamy, S. Balamurali, and S. Kadry, "Diagnosis of diabetic retinopathy using multi-level set segmentation algorithm with feature extraction using SVM with selective features," *Multimed. Tools Appl.*, 2019.
- [4] J. Krause *et al.*, "Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy," *Ophthalmology*, vol. 125, no. 8, pp. 1264–1272, 2018, doi: 10.1016/j.ophtha.2018.01.034.
- [5] T. Araújo *et al.*, "DR | GRADUATE: Uncertainty-aware deep learning-based diabetic retinopathy grading in eye fundus images R Microaneurysms Hemorrhages Hard exudates," *Comput. Sci. Eng. Med. Med. Image Anal.*, vol. 63, 2020, doi: 10.1016/j.media.2020.101715.
- [6] I. Qureshi, J. Ma, and Q. Abbas, *Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning*. Multimedia Tools and Applications, 2021.
- [7] X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, and T. Wang, "Convolutional Neural Networks Based Transfer Learning for Diabetic Retinopathy Fundus Image Classification," *10th Int. Congr. Image Signal Process. Biomed. Eng. Informatics*, no. 978, 2017.
- [8] S. Lahmiri and A. Shmuel, "Variational mode decomposition-based approach for accurate classification of color fundus images with haemorrhages Extract first Compute texture," *Opt. Laser Technol.*, 2017, doi: 10.1016/j.optlastec.2017.05.012.
- [9] D. J. Hemanth, O. Deperioglu, and U. Kose, "An enhanced diabetic retinopathy detection and classification approach using deep

- convolutional neural network," *Neural Comput. Appl.*, vol. 0, 2019, doi: 10.1007/s00521-018-03974-0.
- [10] K. K. Palavalasa and B. Sambaturu, "Automatic Diabetic Retinopathy Detection Using Digital Image Processing," *2018 Int. Conf. Commun. Signal Process.*, pp. 72–76, 2018.
- [11] X. Luo *et al.*, "MVDRNet: Multi-view diabetic retinopathy detection by combining DCNNs and attention mechanisms," *Pattern Recognit.*, vol. 120, p. 108104, 2021, doi: 10.1016/j.patcog.2021.108104.
- [12] S. Pao, H. Lin, K. Chien, M. Tai, J. Chen, and G. Lin, "Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network," *J. Ophthalmol.*, vol. 2020, 2020.
- [13] P. Costa, A. Galdran, and A. Smailagic, "A Weakly-Supervised Framework for Interpretable Diabetic Retinopathy Detection on Retinal Images," *Adv. Signal Process. Methods Med. Imaging*, vol. 3536, no. JULY 2017, pp. 1–11, 2018, doi: 10.1109/ACCESS.2018.2816003.
- [14] "Diabetic Retinopathy Detection | Kaggle." <https://www.kaggle.com/c/diabetic-retinopathy-detection> (accessed Dec. 22, 2021).
- [15] A.KALAISELVI, NAGARATHINAM, T. D. PAUL, and M.ALAGUMEENAAKSHI, "DETECTION OF AUTISM SPECTRUM DISORDER," *Turkish J. Physiother. Rehabil.*, vol. 32, no. 2, pp. 926–933, 2020.
- [16] A. Bin Tufail *et al.*, "Diagnosis of Diabetic Retinopathy through Retinal Fundus Images and 3D Convolutional Neural Networks with Limited Number of Samples," *Hindawi Wirel. Commun. Mob. Comput.*, vol. 2021, 2021.
- [17] R. Casado-Vara, A. M. del Rey, D. Pérez-Palau, L. De-La-fuente-valentín, and J. M. Corchado, "Article web traffic time series forecasting using LSTM neural networks with distributed asynchronous training," *Mathematics*, vol. 9, no. 4, pp. 1–22, 2021, doi: 10.3390/math9040421.
- [18] P. N. Srinivasu, J. G. Sivasai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with mobile net v2 and lstm," *Sensors*, vol. 21, no. 8, pp. 1–27, 2021, doi: 10.3390/s21082852.

TELEFIT – An IoT enabled health care Device

Soma Prathiba

Department of Information technology
Sri Sairam Engineering College
Chennai,India
prathibha.it@sairam.edu.in

Srinithi S

Department of Information technology
Sri Sairam Engineering College
Chennai,India
sec20it013@sairamtap.edu.in

Shyamkumar M

Department of Electronics and
Communication
Sri Sairam Engineering College
Chennai,India
sec19ec048@sairamtap.edu.in

Kirthiga M

Department of Electronics and
Communication
Sri Sairam Engineering College
Chennai,India
sec19ec062@sairamtap.edu.in

Pravin Kumar P

Department of Information technology
Sri Sairam Engineering College
Chennai,India
sec20it076@sairamtap.edu.in

Yashwanth Krishnan B

Department of Information technology
Sri Sairam Engineering College
Chennai,India
sec20it126@sairamtap.edu.in

Shabika Fathima M

Department of Information technology
Sri Sairam Engineering College
Chennai,India
sec20it074@sairamtap.edu.in

Abstract— Amidst this uncertain situation of the Covid pandemic happening worldwide, It is necessary to protect the health workers too. Since they are more prone to the dangerous virus than any other. To reduce the risk of virus infection spreading to Health care workers, this paper proposes the concept of wearing a smart Healthwear band, an IoT monitoring device integrated with a mobile application for monitoring the patients with a contact-less mechanism. Healthware is placed on the Covid infected patients which sense the temperature, pulse, and oxygen level of the patient. This was monitored manually before with the equipment present in the hospitals. This Health wear device automatically collects the data of patients and alerts the Nurses and doctors in case of emergencies. This reduces the contact of Healthcare workers with infected patients. This paper briefly describes the working of the IoT enabled healthcare device and the methodology used in detail. The cost of the proposed device is very less compared to other existing systems.

Keywords— Health monitoring device, Mobile application, and contact free heath wear.

I. INTRODUCTION

Internet of Things has now become the most important technology of the 21st century. IoT is an interrelated internet-based system that connects, Collects, and transfers data over devices wireless and without any sort of Human Interruption. It has made human beings' lifestyles easier as every device can be connected using it. Also, IoT technology has been consistently evolving and new connectivity strategies are being found then and there since its Invention. This rather impacts Industries of all kinds and our Personal life as well. [1] Penetration of the Internet of things in the Health sector is expected to hit around \$409.9 billion by 2022. Biosensors, which work on the principle of IoT, have a major impact on the digitalization of the Health sector. It helps in collecting data from the sensors and transmitting it into a web application or a mobile application. In our proposed work, we are using various types of Biosensors and implementing the concept of IoT to collect the patient's Pulse rate, Body temperature, and oxygen count in the blood. It is stored in a database and can be accessed

when required. Also, there is an alert system that is required when a patient is in a medical emergency such as massive fluctuations in the pulse rate or oxygen level in the patient's blood going down drastically and he/she needs immediate physical assistance. IoT is involved in the proposed theory to cover a larger patient population to reduce the requirement for more healthcare workers. It is also used to give a warning or alert in case of medical emergencies to avoid frequent contact between patients and healthcare workers. Concepts of IoT used in the proposed theory are Sensors, connectivity, data processing, and user interference

Using the concepts of Information technology and Telenursing in Telenursing it provides nursing services in healthcare activities. This method reduces the direct interaction between the nurses and patients. Then how would they interact with each other? The nurse monitors the health condition of the patients through a telecommunication device. There are many examples of Telenursing. Starting from the phone we use, interactive video, phone triage, internet support, and remote telemonitoring are some of the examples through which telenursing is performed practically.

Telenursing extends the nursing service to more patients. Let us consider the practical example of real-life nurses. A nurse can reach and care for only 5 to 7 patients through direct contact in a home-care delivery. Whereas in the case of the Telenursing concept, a nurse has to check more patients at the same time. Considering the major components used in telenursing, hardware (physical computer body), computers (laptops and desktops), palm Computers, telephones, graphic cards, Internal Memory (RAM), External Memory (Flash Device) The nurse can acquire the following informatics through our device, The nurse can collect the current stage of patients' health data from patients, The collected data can be accessed using knowledge, All the retrieved data is been documented. Critical thinking skills is been utilized for predicting the output, Nursing interventions is been provided, It continues to monitor the health of the patients by

utilizing the available technology Implications of telenursing from the patient side helps to monitor the changes daily and informs the care specialist by providing opportunities for early intervention. Unless there is an emergency need, the intervention between doctors and patients is reduced completely along with the visit to or by providers. It also provides patients with the education of telenursing and telecommunications. It removes geographic barriers. Nursing can be provided to remote locations. Implications of telenursing from the provider (i.e.) doctors' side increase the competencies and scope of the practice. Focus more on client safety and guided monitoring systems. It allows easy client decision-making in case of any need. The implementation of telenursing creates a professional practice environment. It tends to provide expert and specialized care from staff and caretakers. Joint decisions are encouraged during consulting with staff even at far places. Reduces the risk of infection spreading to direct methods and thus enables a safer environment.

Continuous observing of patients are the part and parcel of the lives of the health issues. In recent times it has become very common. The health workers has to keep in track of all their patients so as to prevent any unnecessary negative incidents. In order to reduce the burden on the health workers far-off and safe guard the health workers as well as the patient an IoT based health monitor system along with a mobile application is proposed in our paper. The IoT health monitor system allows for remotely monitoring of multiple patients and the health worker can monitor it through a mobile application developed using android Studio. TELEFIT monitors patient pulse, oxygen level and temperature of the person by employing a pulse oximeter and temperature sensor. It then transmits this data to the Firebase (cloud) using the Wi-Fi module or the Nodemcu. The TELEFIT Health wear is placed near the patient and the necessary values are transmitted through the cloud. This allows the doctor to monitor the patients remotely without the danger of infection. One doctor can monitor more than 500 patients at a time. The doctor gets an alert just in case of health fluctuations or emergency. This allows to the doctors to take appropriate care and make the patients feel secured as the patients are monitored continuously. The next section gives a brief information about the health monitoring devices that are already available in the market and the major advantage of our TELEFIT.

II. LITERATURE SURVEY

In recent times many projects and devices are developed related to Covid as mentioned in [8]. Already the count of people who need monitoring and assistance is increasing year by year. To assist the people without involving the doctors all the time here in Table 1. Summarizes a few features of the products which are already available with our proposed system. A short description of the budget required for making similar projects is also mentioned below.

TABLE I. DESCRIPTION OF EXISTING SYSTEMS

TITLE	DESCRIPTION	COST	DRAWBACK
Telehealth monitoring system in Rural areas [2]	Sends the data like pulse, oxygen level in blood, ECG signals etc. of the patient to cloud	Initial setup cost for devices for taking the pulse, oxygen level, ECG signal etc. Cloud database cost for deployment	The pulse and oxygen level cannot be taken by the patient by himself. Needs someone to take the readings and send to the cloud
An approach to make way for intelligent ambulance using IoT [3]	The system will monitor and measure vital signs and send the data to the cloud database in an exceedingly PHC where doctors will monitor the information and be ready to incorporate an ambulance in an emergency.	Expenses for giving pay for the ambulance driver Cloud database cost for deployment Hardware expenses	Not user friendly Requires someone to check in the backend always
Autonomous monitoring system[4][5][6]	Our study aims at developing an automatic system for non-contact measurement of important body parameters.	Uses cameras in addition to sensors. Hence cost is more.	
Traditional thermometer and pulse oximeter		Thermometer- 7\$ Pulse oximeter- 180\$ Total 187\$	The doctor cannot view the values of the pulse and oxygen level through a app The doctor or nurse should make contact with the patient in order to observe the condition of the patients

The necessary conditions to develop a wearable health device were quoted in [7]. It should be comfortable for both the user to wear and the physician to use them. The paper highlights the compatibility of the device and stresses that the device has to be lightweight. In a survey [9] the number of monitoring devices for elderly people has reached its peak. This pandemic has boosted the count multifold. Also in [10], it is mentioned that these technologies are in phase and will surely reach the market at a bigger level in forthcoming years. Last decade as quoted in [11], real-time health monitoring devices have drawn the attention of the research people. Making sounds that the health monitoring device will be on the top need, in [12] a fully automated health monitoring in the home was done. Skin-based heath wear has even more great potentials to revolutionize the use and need of the devices. Wearable, wireless and low-cost applications are needed as in [13]. One such highly recommended device is thus developed, summing up all the advantages and reducing the disadvantages. As mentioned in [14], [15], and [16] the number of advantages of health monitoring devices

is enormous. Few of them are easy diagnosis, data collection, Cost-efficient, affordable, and early detection of diseases. Considering the advantages we have developed TELEFIT, a wearable device that helps to prevent the spread of disease to the health workers. The features of TELEFIT is described in detail in the next section.

III. PROPOSED SYSTEM

As quoted in the last sessions the safety for health workers is regretted most of the time in hospital. Thus the proposed system is developed in such a way that the patients can be monitored without any contact. The contact less monitoring of the patients is carried out with two major segments. The first segment is the hardware device which consists of a pulse oximeter integrated with the temperature sensor. The second segment is the app module through which the parameters can be monitored. The cloud acts as an interface between the first segment and the second segment. The firebase real-time database is used for data storage and implementation of the system.

A. The hardware device

The hardware health wear device is responsible for data acquisition from the patient's end and transfers it to the firebase real-time database. It is placed in the patient's end for automatically monitoring the data such as the oxygen level, pulse count, and the temperature of the person without any contact of the health workers and the patients. This device is mainly based on circuitry and it plays a major role in transmission of the data from the patients to the doctor. The components of the TELEFIT are mentioned in table 1. The Nodemcu or the Wi-Fi module acts as both the microcontroller and the transmitter which transmits the data that has to be monitored to the cloud. The MLX90614 sensor is the infrared temperature sensor that determines the temperature of the person without any contact. The pulse oximeter sensor is a biometric-based sensor, which helps to get the oxygen level and pulse count of the patients. Using the principle of photoplethysmography, the MAX30100 does all the sensing by utilizing its internal LEDs to bounce light off the arteries and arterioles inpatient finger's subcutaneous layer and sensing how much light is absorbed with its photo detectors. This data is passed onto and analyzed by the MAX32664 which applies its algorithms to determine heart rate and blood oxygen saturation (SpO₂). SpO₂ results are reported as the percentage of hemoglobin that is saturated with oxygen. The connections are made according to the circuit diagram shown in Fig 1. The connections are done with connecting wires or the jumper wires. To secure the connection the wires are soldered and insulated using insulation tapes. The nodemcu or the Wi-Fi module is powered with a battery of 9V.

TABLE II. MAJOR COMPONENTS OF TELEFIT-HARWARE.

S.NO	COMPONENT NAME	SPECIFICATION
1	Temperature sensor	MLX90614
2	Pulse oximeter	MAX30100
3	Nodemcu/micro controller	ESP8266

The Arduino IDE is used to provide instructions to the microcontroller. D1, D2, D3, and D4 which are marked in the Nodemcu are the data pins. The pins D3 and D4 are used to get the oxygen level and pulse count of the patient as they are connected to the pulse oximeter. The pins D2 and D1 which are connected to the MLX90614 (temperature sensor) are to get the temperature of the patient. 3V refers to the power pin, this is used to get the power for the sensors. GND pin refers to the ground pin. The values are read from the data pins and are saved in the firebase real-time database as shown in fig () . To save the data in the firebase real-time database as shown in Fig (), the nodemcu and the firebase console are connected using the credentials. As a real-time database is used the values keep on updating automatically as the sensor values change. The updated values can be monitored using the Mobile application which is integrated with the telefit. The flowchart of the TELEFIT is shown in Fig 2.

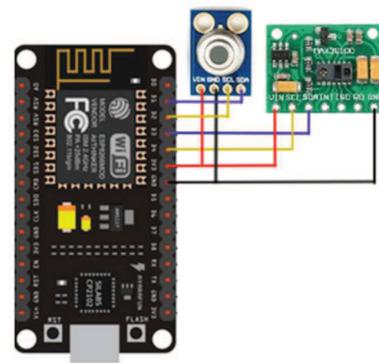


Fig. 1 Circuit diagram of the health ware device.

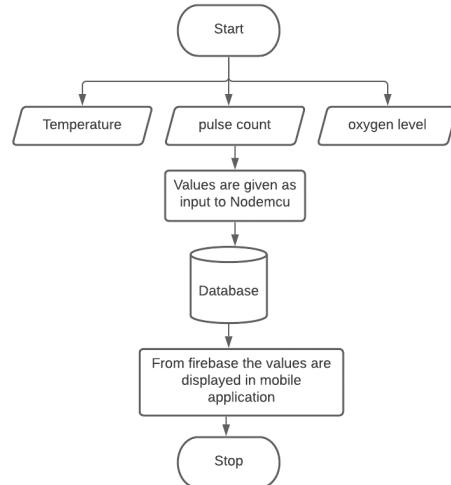


Fig. 2 Flowchart of TELEFIT, a health wear.

B. The mobile application

This is the segment which the health workers use to monitor the pulse count, oxygen level, and temperature of the patient. This mobile application is developed with android studio and the app is connected to the firebase. The mobile application has a dashboard and a page for displaying the data of the patient as shown in Fig 3. The dashboard has the details of

the bed number. For each bed number, the data such as the oxygen level, pulse count, and the temperature of the patient are displayed. The data is obtained from the real-time database as in Fig 4 and updated on the screen as soon as the value is updated in the cloud. This application helps the health worker to monitor the patients without any contact. This also allows long-distance monitoring making it more advantageous.



Fig. 3 Dashboard

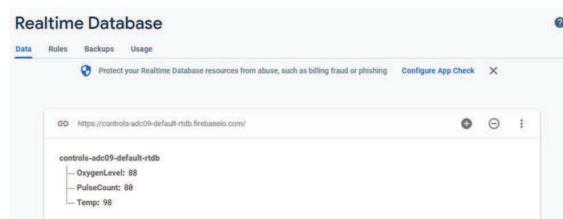


Fig. 4 Real-Time database

C. Role of TELEFIT in reducing the disease spread:

The proposed wearable health device, the TELEFIT can be placed in the wards where many patients need to be monitored, it can be used by users whose real-time data has to be monitored. The doctor or the health worker can check their patient's data right from their inwards. Timely check of the patients is now automated, thus making them little safer. In case of an emergency, the doctor can view it and reach the specific patient well in advance, as the values are being updated. As the device is connected to the mobile application of the doctor, the values such as temperature, pulse count and oxygen level of the patient is updated to the

firebase. This enables the doctor to monitor their patients without any contact. When the patients keeps his/her index finger on the device the values are taken to firebase via the nodemcu. This makes the contactless data transfer possible making the spread of disease nullified. The results obtained with this health wear is detailed in next section.

IV. RESULT & DISCUSSIONS

The process of monitoring the patients without any contact is done using telefit. The values obtained through the device is shown in Fig 5. The concept of distant monitoring of patient health conditions by Healthcare workers is an absolute requirement in the present COVID pandemic situation. A few of the important points are mentioned below

1) Continuous monitoring possible: Usually monitoring of Vitals is possible only during rounds for patients inwards. But with TELEFIT, we can monitor all the time.

2) Distant monitoring possible:

Though regular continuous monitoring is done for ICU patients, personnel should be at a reachable distance to look into monitors or listen to alarms. It can be done from a very distant place with TELEFIT.

3) Periodic alerts of changing conditions:

As there won't be any attendant accompanying the patient in COVID wards, he has to call for help or to complain about any symptom. With TELEFIT patient can assure himself that someone is looking at his condition and any changes in his Vitals are conveyed to Healthcare personnel timely.

4) Minimum exposure of Healthcare personnel

TELEFIT decreases the number of times Healthcare personnel has to get into direct contact with the patient

5) Tool for monitoring during teleconsultations

Home treatment with regular teleconsultations has become the modality of outpatient management. Such devices are of great help for follow up

With many such applications in a Healthcare setup, health care devices like TELEFIT are of great help in the effective management of many health-related problems and diseases with ease of monitoring and follow-up.

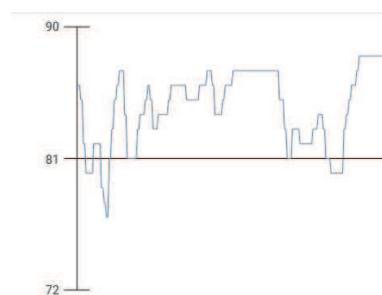


Fig. 5 Output graph of pulse (beats/minute)

V. CONCLUSION

This paper focuses on monitoring the pulse count, temperature, and oxygen count of the patient using the sensors to detect the fluctuations and a mobile application that is structured with the help of the android studio, both being connected to the firebase. With the help of this, Treatment for COVID patients can be achieved easily even amidst the prolonging pandemic along with ensuring safety for healthcare workers. In the hardware segment, various sensors are used to monitor the above-mentioned parameters and the readings are streamed directly to the firebase using nodemcu. If any fluctuations in the reading are detected the mobile application which is again integrated with nodemcu sends an alert. Thus the patient's health is monitored without contact in person in real-time more easily and efficiently. The main goal of this system is to reduce the risk of contact between doctors or Health line workers and patients. This system provides accurate and error-free results with the help of well-designed components incorporated in the system. Thus there is nil chance of having fake alerts or misleading results. Also, it is more economical than other devices which are available in the market. The main key factor of the success of this device will be the efficiency of data transmission. Nodemcu was involved in data transmission up to 4.5Mbps. Thus the live status of the patient could be monitored from anywhere anytime. This system also involves medical status history storage which helps in the time of consultations and also in emergencies. The data stored in the database are secured and security screenings are done for the user before accessing it, thereby maintaining their privacy. This reduces the strength of data leaks and any other kind of misuse. Despite having all major advantages, it has some common disadvantages like miscommunication due to lack of maintenance and also due to abnormal medical conditions. Thus on the whole it is a life-saving and comprehensive system in the field of telenursing, thereby assisting medical practitioners.

REFERENCES

- [1] M. A. Akkaş, R. Sokullu, and H. E. Çetin, 2020. Healthcare and patient monitoring using IoT. *Internet of Things*, 11, p.100173.
- [2] Biradar, Vijayalaxmi, and G. Durga Sukumar. "Tele Health Monitoring System in Rural Areas Through Primary Health Center Using IOT for Covid-19." *The Fusion of Internet of Things, Artificial Intelligence, and Cloud Computing in Health Care*. Springer, Cham, 2021. 157-173.
- [3] Venkatesh, H., Shrivatsa D. Perur, and M. C. Jagadish. "An approach to make way for intelligent ambulance using IoT." *International Journal of Electrical and Electronics Research* 3.1 (2015): 218-223.
- [4] Kim, Jong-Hoon, et al. "SPAMMS: A sensor-based pipeline autonomous monitoring and maintenance system." *2010 Second International Conference on COMmunication Systems and NETworks (COMSNETS 2010)*. IEEE, 2010.
- [5] Kang, Dongho, and Young-Jin Cha. "Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging." *Computer-Aided Civil and Infrastructure Engineering* 33.10 (2018): 885-902.
- [6] Rice, Jennifer A., et al. "Flexible smart sensor framework for autonomous structural health monitoring." *Smart structures and Systems* 6.5_6 (2010): 423-438.
- [7] Ha, M., Lim, S. and Ko, H., 2018. Wearable and flexible sensors for user-interactive health-monitoring devices. *Journal of Materials Chemistry B*, 6(24), pp.4043-4064.
- [8] Abdullah, A., Ismael, A., Rashid, A., Abou-ElNour, A. and Tarique, M., 2015. Real time wireless health monitoring application using mobile devices. *International Journal of Computer Networks & Communications (IJCNC)*, 7(3), pp.13-30.
- [9] Durán-Vega, L.A., Santana-Mancilla, P.C., Buenrostro-Mariscal, R., Contreras-Castillo, J., Anido-Rifón, L.E., García-Ruiz, M.A., Montesinos-López, O.A. and Estrada-González, F., 2019. An IoT system for remote health monitoring in elderly adults through a wearable device and mobile application. *Geriatrics*, 4(2), p.34.
- [10] Korhonen, I., Parkka, J. and Van Gils, M., 2003. Health monitoring in the home of the future. *IEEE Engineering in medicine and biology magazine*, 22(3), pp.66-73.
- [11] Abdullah, A., Ismael, A., Rashid, A., Abou-ElNour, A. and Tarique, M., 2015. Real time wireless health monitoring application using mobile devices. *International Journal of Computer Networks & Communications (IJCNC)*, 7(3), pp.13-30.
- [12] Tamura, T., Ogawa, M., Yoda, M. and Togawa, T., 1998. Fully automated health monitoring system in the home. *IEEJ Transactions on Electronics, Information and Systems*, 118(7-8), pp.993-998.
- [13] Honda, W., Harada, S., Arie, T., Akita, S. and Takei, K., 2014. Wearable, human-interactive, health-monitoring, wireless devices fabricated by macroscale printing techniques. *Advanced Functional Materials*, 24(22), pp.3299-3304.
- [14] Walker, R.K., Hickey, A.M. and Freedson, P.S., 2016. Advantages and Limitations of Wearable Activity Trackers: Considerations for Patients and Clinicians. *Clinical journal of oncology nursing*, 20(6).
- [15] Pandey, G. and Vora, A., 2019. Open electronics for medical devices: state-of-art and unique advantages. *Electronics*, 8(11), p.1256.
- [16] Agarwal, V., 2018. Smart sensors for structural health monitoring-overview, challenges and advantages. *Sensors & Transducers*, 221(3), pp.1-8.

IoT Based Collaborative Industrial Gas Prediction and Warning Using Machine Learning

B.R.Sathishkumar

Associate Professor

*Department of Electronics and Communication Engineering
Sri Ramakrishna Engineering College
Coimbatore, India
sathishkumar.b@srec.ac.in*

Sree Dhviya M

Student

*Department of Electronics and Communication Engineering
Sri Ramakrishna Engineering College
Coimbatore, India
msdhviya4@gmail.com*

S.Allirani

Associate Professor

*Department of Electrical and Electronics Engineering
Sri Ramakrishna Engineering College
Coimbatore, India
allirani.saminathan@srec.ac.in*

Madhumithaa V

Student

*Department of Electrical and Electronics Engineering Sri
Ramakrishna Engineering College
Coimbatore, India
madhumithaavl@gmail.com*

Abstract

Air pollution is a major environmental issue and it affects the health of humans. A major cause for this is industries and due to this, the workers are affected to a great extent. To take precautionary measures an idea is proposed to detect the harmful gases in the industry surroundings. Anyone can get the level of the gases from anywhere using their mobile phone using this proposed solution. This is done using the Internet of Things. The sensor data is collected and it is continually plotted for each second as a graph to monitor the level of gas. A messaging system is used for the transmission of the sensor data by integrating it with a GSM module using a SIM card. By doing this one can ensure that the data are being processed as soon as possible. This will prove to be one of the methods of transmission of data to other electronic mobile sources within a very short period about the environmental conditions. There is also an integration of machine learning into this system which is used to predict pollution levels. With the data set provided, the machine learning algorithm calculates the gas levels which can be used to predict the gas levels in the future. This will prove to be very helpful in taking precautionary measures for air pollution and to reduce the upcoming problems caused due to pollution and this can lead to a cleaner and greener environment.

Keywords: Pollution, Messaging, Internet of Things, Machine Learning, Gas Sensors.

I. Introduction

The recent trends are about driving towards a clean environment. A clean environment means reducing the number of toxic substances in the environment. Among various types of pollution, the major affecter of our environment is air pollution. Air pollution is caused due to many factors and it has adverse effects. The pollution levels can be reduced only by reducing the number of pollutants entering the atmosphere. But one has no data on the number of gases emitted from each source and which pollutant affects the most [1]. So, an idea has been proposed to monitor gases in the industry to find and warn us about the level of harmful gases present in the surrounding. By detecting the level of gases, one might be able to find its source and eradicate it in some other cases where the high gas levels may cause serious health injuries, people in that environment can be alerted. As there are many mixed gases in industries and their surroundings, the integration of two types of gas sensors have been done to get data on various gases. It can further be integrated

with other gas sensors according to the need of the industry. The average safety level of gases can be calculated from the previously available data and the limit of harmful gases may be fixed and the threshold value for the gases for warning may be set. The major setbacks of the recently available gas monitoring systems^[3] are that they require manual monitoring all the time through the main control systems. So, to do this, a messaging system can be inserted for the transmission of the sensor data by integrating it into a GSM module using a SIM card. By doing this one can ensure that the data are being processed as soon as possible. This will prove to be one of the fastest methods of transmission of data to other electronic mobile sources with a very short period in respect to environmental conditions. Manual methods do not prove to be efficient and accurate and there might be great losses when they are not monitored properly. Thus, this system will be the quickest means of alert system with high efficiency.

II. Related Work

The main idea of the system is to prevent the environment from getting polluted and to maintain the environment in the cleaner way possible. Pollution is a major threat to the environment and also to the beings of the environment. In the existing system the air pollution level is continuously monitored and the system that was proposed was to make the existing system of air pollution level monitoring more feasible and to reduce the cost of the system and at the same time, it is seen the efficiency of the system is increased [1]. In the existing system the data are fetched through the sensor and are stored in the cloud for future processes. The data that is fetched from the existing system can be used for various purposes. The proposed system uses various kinds of sensors and it is used to monitor the air pollution level in homes. The data that are obtained are displayed on the Raspberry Pi web server. The data that is received from the sensors are stored and the mean value of the data is displayed on the webserver. The data obtained are first stored in the cloud and at the end, the data are transferred to the webserver. The system is designed to give awareness to the

people about the level of pollution in the surrounding. The existing system is mostly acceptable for home usage. The systems that exist are kept on detection of the gas concentration of the gas level in the surroundings. The system alerts the user and so the necessary precautionary measure can be taken.

III. Methodology

The microcontroller used here is Raspberry Pi and it acts as the CPU controlling all functions. The gas sensors used are analog sensors and the data set obtained will be analog which is difficult for further processes such as plotting graphs hence Analog to Digital converters are used and the converted data is fed to the cloud from the microcontroller. The cloud data is displayed on mobile phones whose numbers are given as input to the microcontroller. The machine learning algorithms are designed according to the sensors and the module is run in the shell for which the input is given as units on time in minutes. The GSM module is which can be used to send messages in case of emergency wherein it should contain a SIM card of its own.

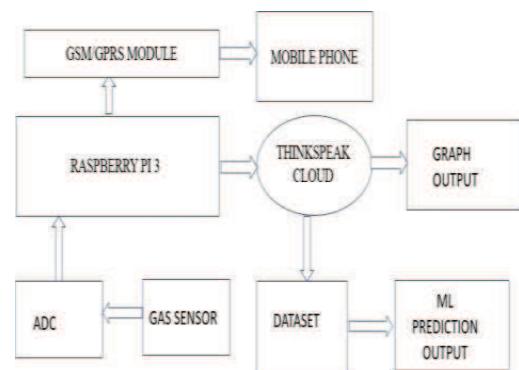


Fig.1 Block Diagram

This allows the Raspberry Pi to generate a message when the previously set threshold value for gases exceeded and send it to the GSM module. Finally, the message will be sent to the corresponding mobile phones or other electronic devices with specified numbers. The Raspberry Pi 3 B, The MCP3008 8- channel 10-bit analog to digital converter, GSM/GPRS MODEL 8001, Grove - Gas Sensor (MQ2), Gas Sensor-(MQ7) are the Hardware Components used in the Proposed

System as mentioned in Fig. 1.

a. RASPBIAN

The Raspbian is a Debian-based PC working framework for Raspberry Pi. Raspberry Pi line's low-execution ARM CPUs [6]. Raspbian is an OS for the Raspberry pi 3 microcontroller. Raspberry Pi is a single-board computer.

b. THINGSPEAK

ThingSpeak is a platform used for the Internet of Things. In this platform, one can store and access the sensor data in the cloud and develop an IoT application. ThinkSpeak provides the framework to develop IoT-based applications. The ThingSpeak helps to analyze and visualize data using their application. The data collected from Sensor are sent to ThingSpeak via Raspberry Pi. ThingSpeak is an API to store and collect the data from the sensor using the HTTP protocol over the internet [2]. ThingSpeak can be viewed in any electronic gadget using the application. ThingSpeak is the software tool used for implementing the proposed work.

IV. Proposed Work

The Raspberry Pi 3 B, The MCP3008 8-channel 10-bit analog to digital converter, GSM/GPRS MODEL 8001, Grove - Gas Sensor (MQ2), Gas Sensor-(MQ7) are the Hardware Components used in the Proposed System. To obtain precise results, linear regression is found to be the best suitable for this dataset. The attribute x and the attribute y are the input and output variables respectively. The variable y is to be predicted. In figure 2 the relationship

X versus Y

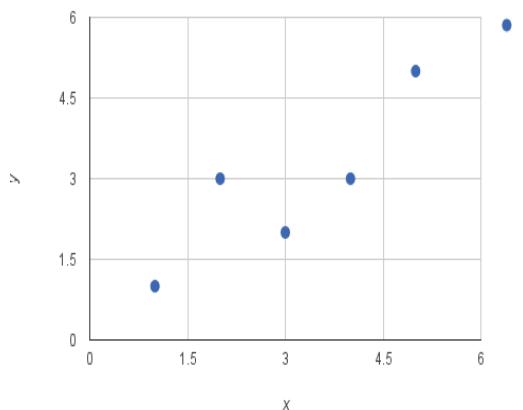


Fig. 2. A plot of Simple Linear Regression

between x and y is shown. The graph between x

and y is almost linear. As in, draw a line somewhere diagonally from the bottom left of the plot to the top right to generally describe the relationship between the data [3].

General equation of Linear Regression

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Where

Y_i - Dependent Variable

ε_i - Random Error term

β_0 - Population Y-intercept

β_1 - Population Slope Coefficient

X_i - Independent Variable

The x_i and y_i are the i^{th} value in the dataset and i refer to the value of x or y. The Y is found separately for each gas using a separate value from the dataset.

Estimating Error

Calculate an error for our predictions called the Root Mean Squared Error or

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\text{predicted}_i - \text{actual}_i)^2}{N}}$$

V. Experimental Results

A. Output

The hardware setup is made by assembling all the components as mentioned in the block diagram as shown in Fig. 3. The Raspbian OS is installed in the Raspberry Pi. Then the gas sensors MQ-2 and MQ-7 are connected with their analog pins to the Analog to Digital Converters. Then the ADC is connected to Raspberry Pi. The code is executed. For machine learning, the code is executed in the VNC viewer and the result is obtained [4]. Then the GSM module is integrated into the Raspberry Pi after inserting a SIM card into the GSM module. All the connections are made using jumper wires and a 12V power supply is given to the GSM module apart from the power supply to the Raspberry Pi [5-6].

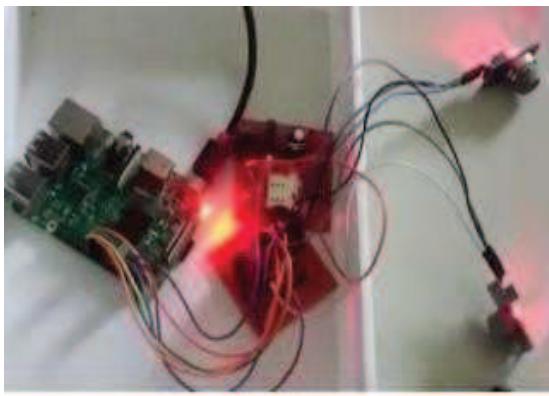


Fig 3. Experimental Hardware setup

B. Output for Gas Level Detection as Graphs

The gas sensors are connected to an analog to digital converter and converted into digital data and then to the control system, that is the Raspberry Pi, the data from the sensors is transmitted to the cloud. The cloud data is then further processed and plotted in the form of graphs. Through this one can monitor the level of gases in the surroundings.

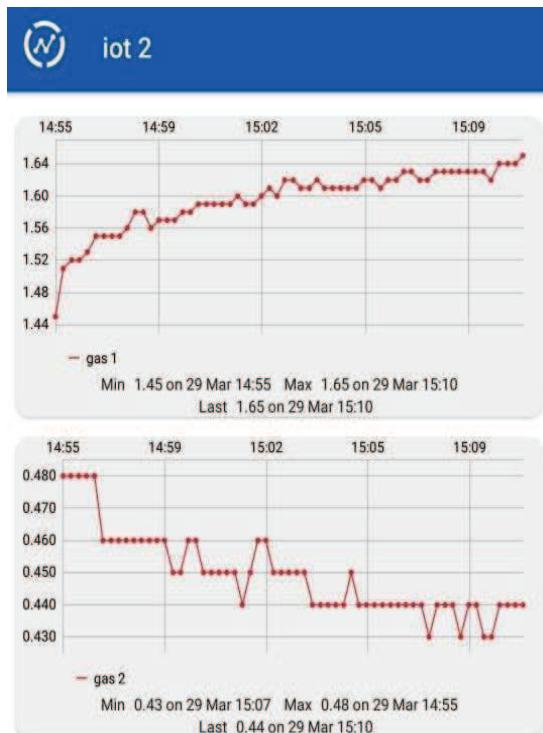


Fig.4 Graph Output in Mobile Phone

The greatest advantage of this system is that the graphs plotted can be made available to

be viewed in any electronic gadget that can install this application. The graphs so obtained are depicted in the test data given below in Fig. 4.

C. Output for Machine Learning

The machine learning algorithm is run to determine the content of gas in that surrounding by giving input of the time in minutes. As expected, the output for the level of gases at a given period is calculated using the machine learning algorithms and the values are displayed in the output screen of the system. The obtained results are depicted below in fig. 5.

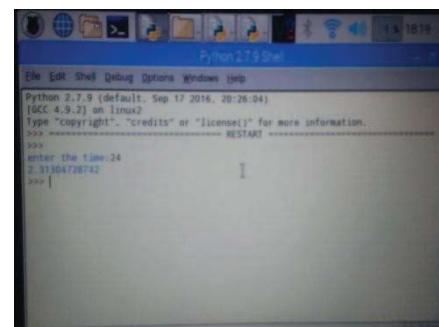


Fig. 5 Machine learning Prediction

Here, the example is taken for 24 minutes the output value shows the predicted value of gas after 24 minutes.

D. Output for Alert System

The major idea of this project is the implementation of a rapid alert system with high accuracy and efficiency. The alert system is induced to send a message with a warning note as soon as the level of gas in the surroundings increases beyond the threshold value.

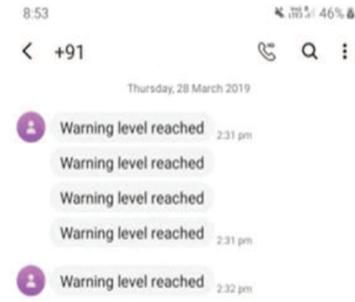


Fig. 6 Warning Message Screenshot

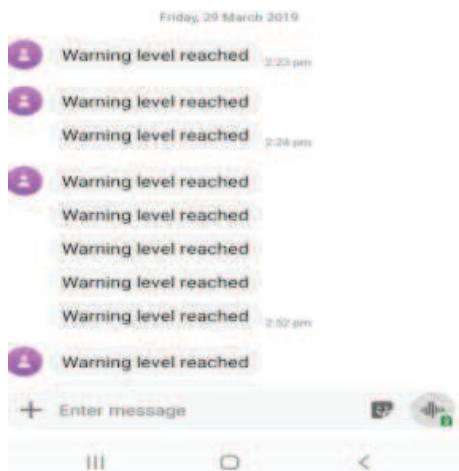


Fig. 7 Warning Message Screenshot

The integration of the GSM module enables the system to send the messages immediately as the level hits the threshold value which has been calculated and fed into the program [7]. The sample results obtained when the gas level crossed the threshold value can be viewed on mobile phones. Thus, the message has been sent to the respective mobile phones. Figs 6 & 7 shows an example message sent to the user.

VI. Conclusion

This project helps in providing an enhanced alert system for monitoring industrial atmospheric gases and gives very quick alerts of the toxic gas levels. This system is predominant in its way of providing accessible data to everyone in the field and the number of inputs can be increased according to the number of users. Since this project is based on the prediction system of the future pollution level, the necessary preventive measures required can be taken.

References:

- [1] Gagan Parmar, Manju K Chatopadhyay, Sagar Lakhani “An IoT Based Low-Cost Air Pollution Monitoring System”, Proceedings of International Conference on Recent Innovations in Signal Processing and Embedded Systems, pp.524-528, 2017.
- [2] Analyzing MQ-5 Gas Sensor Using ThingSpeak - December 26, 2016.
- [3] Ashutosh Sharma, Poonam Pal, Ritik Gupta, Sanjana Tiwari “IoT Based Air Pollution Monitoring System Using Arduino”, International Research Journal of Engineering and Technology, Volume: 04 Issue: 10, pp.1137- 1140, 2017
- [4] Malaya Ranjan Panigrahy, Riteeka Nayak, Vivek Kumar Rai “IoT based air pollution monitoring system”, Vol-3, Issue-4, pp.571-575, 2017
- [5] Palaghat Yaswanth Sai “An IoT Based Automated Noise and Air Pollution Monitoring System” Vol. 6, Issue 3, pp.419-423,2017
- [6] <https://en.wikipedia.org/wiki/Raspbian>
- [7] Yonghui Xu, Ruotong Meng and Xi Zhao “Research on a Gas Concentration Prediction Algorithm Based on Stacking”, Sensors 2021, vol.21, Issue.05, 2021, <https://doi.org/10.3390/s21051597>.

A Comparative Approach for Investigating an Efficient Method for Reduction of Search Space in Lung Image Analysis

1st Sheetal Pawar

Electronics Engineering

Walchand College of Engineering

Sangli, Maharashtra, India

email: pawarsheetal13@gmail.com

2nd Dr. B.G.Patil

Electronics Engineering

Walchand College of Engineering

Sangli, Maharashtra, India

email:babasaheb.patil@walchandsangli.ac.in

Abstract— Majority of the lung disorders are diagnosed and treated using chest radiographs and volumetric CT scans. The experts are examining mainly the lung parenchyma for disease predictions and treatment. Though the radiology experts could decipher the lung boundary manually, it is very lengthy and susceptible to observer variations. The effective and automated segmentation of the region of interest (ROI) guarantees the speed and accuracy of clinical examinations. This study mainly focuses on investigating an automated and efficient lung segmentation (LS) technique that will effectively reduce the search space and improve the accuracy of clinical investigations. This study reports the comparative analysis of the traditional LS methods viz. segmentation using image processing and Watershed technique with the supervised learning strategy (SLS) using modified U-net. SLS outperformed the existing frameworks and achieves the highest segmentation accuracy of 99.7%.

Keywords—Segmentation, Lung parenchyma, Convolution, Upsampling, SLS.

I. INTRODUCTION

Segmentation inherits the contextual similarity between the different objects residing in an image. The potent segmentation of the medical scans highlights the region of image beneficial in clinical investigations. The majority of the clinical practices like locating tumours, surgery, radiotherapy, evaluation of post-surgery recurrence is assisted with computer-aided segmentation tools. This study mainly focuses on investigating an automated and efficient LS technique that will effectively reduce the search space and improve the accuracy of clinical investigations.

A. Related work

Fig. 1 describes the various techniques employed by researchers for extracting the lung parenchyma from CT images. The segmentation methods can be broadly differentiated based on the techniques used for separating the lung fields from the whole CT image. Thresholding-based (TB) techniques considered inhomogeneity in the intensities of different tissues present in a lung CT. A lung scan is predominantly occupied with air, lung, nodular regions, fat, water, bones, soft tissues, etc. Each of these sub-components has its attenuation values defined in Hounsfield Units (HU). The attenuation range for lungs is -400 HU to -600 HU. Fat, soft tissues, and bones are much denser than the lung region[1]. Early segmentation methods adopted such gray-level variations among parenchyma and its surroundings while thresholding. A CT scan of a patient shows tremendous variations in intensity values on each slice in a CT stack. Hence, it's very difficult to determine a single threshold for LS. Md. Sakif Rahman et al. processed CT images using Otsu's thresholding and followed a border extraction strategy to locate ROI[2]. Lin-Yu Tseng et al. proposed an adaptive TB

method for LS. The lung CT generally consists of a lung region, trachea, airways, and mediastinum wall. They invented a novel solution to remove the large airways and trachea from lung CT. The primitive statistical measures of intensities were used to separate the trachea from the lungs[3]. Such TB methods fail to provide optimum threshold (OT) in CTs with severe pathologic backgrounds.

Morphological operations are widely used to extract information associated with the shape or geometry of an image using a set of non-linear operators. Few researchers combined thresholding and morphological processing to extract lung segments [1][4]. A series of morphological operations like erosion, opening, and closing was followed after thresholding. Such operations enhance the rate of inclusion of lung nodules attached to the mediastinum wall in ROI. But this technique heavily relies on the selection of OT. Few researchers made effective use of RB segmentation techniques. Such techniques compare the two adjacent pixels in an image based on similarities between them. Pixels with similar attributes are grouped in one region. The major RB segmentation methods are region growing, split and merge, convex hull, and Watershed Segmentation (WS). Nihad Mesanovic et al. applied the region growing technique for LS[5]. They initialized the process of region growing by selecting a seed point from a darker region. The split and merge technique follows exactly the opposite procedure than region growing. It splits the original non-uniform image into four quadrants/segments. It merges the neighbouring quadrants depending upon the uniformity of the segments. If any resultant quadrant is not homogeneous it is further divided into four segments. This process is repeated until no further split and merge is possible. Such RB techniques are time-consuming and usually affected by the strategy utilized for growing and merging. Hye Suk Kim et al. eliminated the task of determining OT by combining MP with anisotropic diffusion[6]. Anisotropic diffusion preserved the sharp edges and filtered the noise.

A modified region-based approach of Convex Hull (CH) was used by a few researchers for LS [7]. For a given set of points, CH tries to fit the smallest polygon. F.Liao et al. pre-processed input image with Gaussian filter to remove the distractions. They first extracted a lung mask using thresholding and connected component analysis. The extracted lung component was eroded to separate the left and right parts of the lung. A CH was applied then on the isolated left and right lung components. CH technique helped to include nodules connected to the exterior walls of the lung[8]. Rustin Shojaii et al. proposed a new RB approach called Watershed Segmentation [WS]. Almost all the techniques discussed above are mostly dependent on the gray-level variations in image and thresholding. It's very difficult to compute such thresholds on the scans with severe pathologic

issues. So, advanced techniques were invented. Advanced techniques reduced the dependency of segmentation accuracy on OT. Liang Zhao et al. proposed dense “Convolutional Neural Network (CNN)” for parenchyma segmentation[9]. This technique adopted a 3D fully connected network. They achieved promising results with few parameters. However, training 3D networks is computationally intensive. Hence, such techniques always reduce the input dimensions of an image, leading to the loss of relevant contents. Caixia Liu et al. invented an LS based on Random Forest (RF) [10]. They employed an improved super-pixel generation method for initial segmentation. They invented a novel circle tracing procedure to correct lung contours computed by RF classifier using super-pixels.

Most of the techniques mentioned above created a strong foundation for LS. However, they somehow rely upon the gray-level changes between the pixel intensities and are very time-consuming. The intensity values in lung CT show remarkable variations with changes in acquisition parameters, trans-pulmonary pressure, and volume of different tissue components. Segmentation results using such approaches are not compelling on scans with prominent pathologic issues. Since most of the time, the abnormalities follow the intensity pattern of the internal tissues. So, there is a need to incorporate a robust, accurate, and scalable technique for LS. We inherited the ability of SLS for precise segmentation of the lung fields. We applied a modified 2D U-Net network for LS. The major contributions are

- This paper offers an overview of various approaches employed for parenchyma segmentation.
- This paper will work as a quick start guide for any new researcher interested in biomedical image segmentation.
- We applied a modified 2D U-Net for LS.
- This paper offers a comparative assessment of the widely used LS methods on a common dataset.

The following sections describe the databases used, the metrics computed to evaluate the performance of each technique, brief details of implemented methods, segmentation results, and remarks.

II. DATABASE DESCRIPTION

This comparative study is carried over the two standard databases “Lung Nodule Analysis LUNA16” dataset, derived from “Lung Image Database Consortium / Image Database Resource Initiative (LIDC / IDRI)” and “Lung Nodule Database (LNDb)”. Such databases offer a common platform for the evaluation of the computational techniques developed for lung cancer diagnosis. The intensity values in a CT scan are the function of the adopted acquisition protocol.

A. LNDb:

LNDb is a retrospective dataset of 294 CT scans. A team of 5 radiologists read scans and identified the lesions with nodules having in-plane diameters, $\geq 3\text{mm}$, $\leq 3\text{mm}$, and non-nodule. The database consists of 1897 annotations comprising 1429 unique findings. Nodule information is tabulated in a comma-separated values (.csv) file. The annotation format includes fields representing the scan ID, coordinates for nodule centroid, nodule volume, and corresponding texture class[11].

B. LIDC:

LIDC is a publicly available reference for the evaluation of computational methods in lung cancer diagnosis. It contains 1018 CT scans. Experts followed two-phase annotation process. They annotated lesions in 3 categories based on the size of nodule diameter. The distinct categories are nodules with diameters $\geq 3\text{mm}$, $\leq 3\text{mm}$, and non-nodules[12]. LUNA16 omits the slices whose thickness exceeds 3mm. They discarded the slices with non-uniform spacing. So, finally, the LUNA dataset has only 888 scans. This set of data figured out 36,378 annotations. Most of the radiologists annotated approximately 1186 identical nodules. So, the final set of annotations includes only 1186 nodules. They followed the following annotation format {seriesuid, coordx, coordy, coordz, probability}[13].

III. EVALUATION METRICS

Segmentation of medical scans is an inevitable part of disease diagnosis. We measured the accuracy of segmentation using statistical tools like “Dice Similarity Coefficient (DSC)” and the “Jaccard Index (JI)”. DSC measures the similarity between two sets of images. Consider extracted lung mask (ELM) as the predicted segmentation result and GT as Ground Truth,

$$DSC = \frac{2 * |ELM \cap GT|}{|ELM| + |GT|} \quad (1)$$

$$JI = \frac{|ELM \cap GT|}{|ELM \cup GT|} \quad (2)$$

Both DSC and JI compare the similarity between ELM and GT and assigns a similarity index in the range of 0 to 1 where 1 denotes the highest segmentation accuracy.

IV. IMPLEMENTATION DETAILS

Most of the image processing techniques proposed in the literature made use of MP, TB, and RB-based methods for LS. We implemented and tested the first technique (T-I) for generating lung masks by combining all these image processing strategies. Watershed segmentation is found more accurate among other RB techniques. In this study, we implemented and verified the WS method as technique-II (T-II). We implemented an automated, fast, accurate, and robust parenchyma segmentation technique using modified 2D U-Net for LS. The brief details of all these methods and corresponding results are discussed in the following subsections.

A. Technique-I (T-I)

LUNA dataset has provided the lung CT images along with its GT segmentation result. All the images and corresponding masks are stored in a metadata format. We assume that the axial mid-slice covers the majority of the lung region in CT. The workflow followed for LS using T-I is presented in Fig. 2.

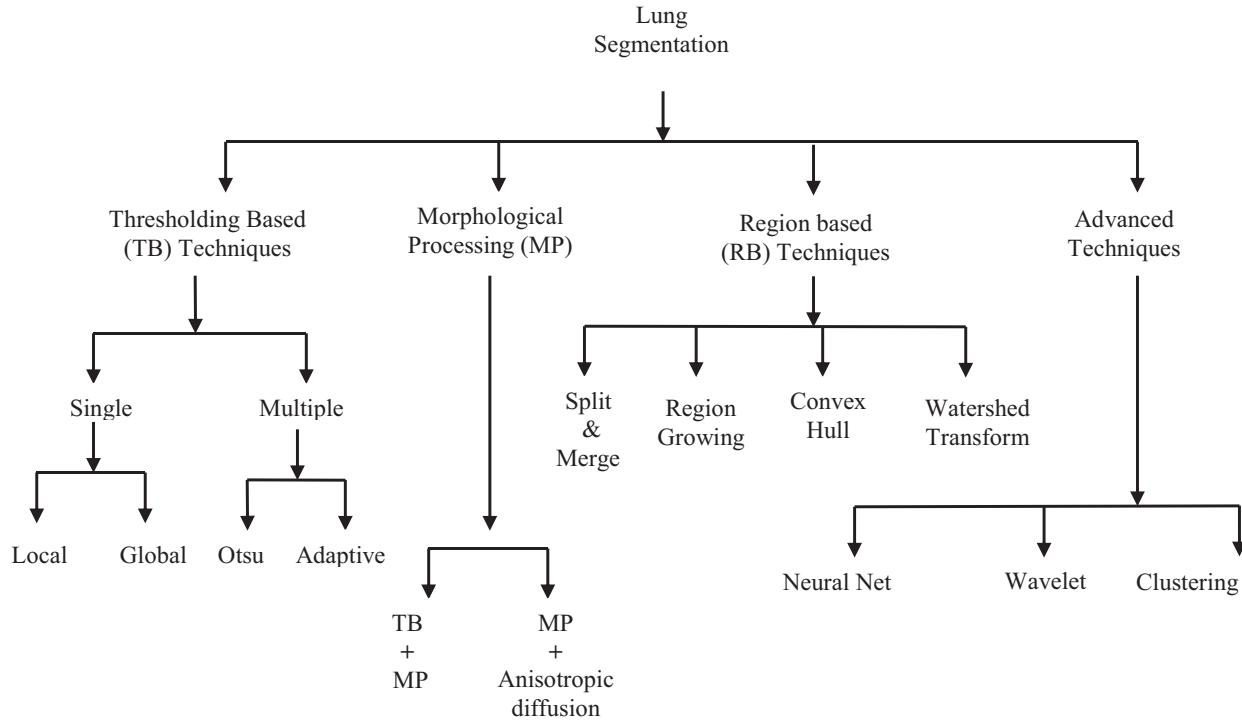


Fig. 1 Lung Segmentation Methods

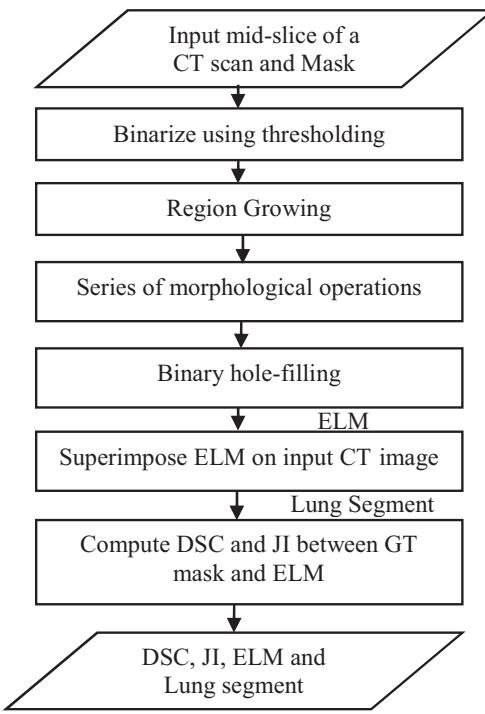


Fig. 2 Workflow of T-I

B. Technique-II (T-II)

The Watershed Segmentation (WS) and its variants are extensively used in segmentation. In WS, the image is considered as a topographic landscape and the pixel intensities project the height of the landscape. In T-II, we applied a marker-based WS approach on the mid-slice of a CT scan. The T-II follows the following subsequent steps for LS as shown in Fig. 3.

C. SLS using modified U-net

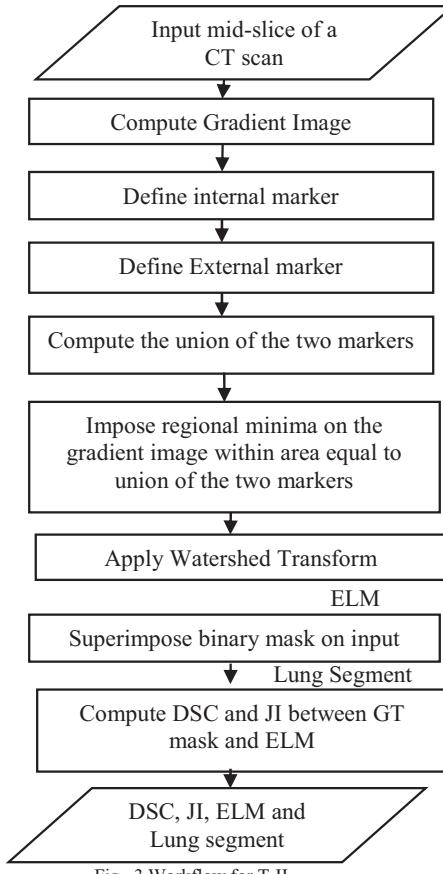
Today, SLS has been accepted as a powerful tool for image analysis. We implemented a novel SLS architecture using modified 2D U-Net as shown in Fig.5. U-Net has been evolved from “Fully Convolutional Neural Network. (FCNN)”. Precise localization and compelling results on small datasets have made U-Net popular in segmentation. U-Net exhibits two symmetric paths namely contractive and expansive. Contractive path assimilates the context and expansive path offers precise localization[14].

Preprocessing of input data

The SLS architecture has been designed to operate on 2D images. Initially, a 3D volume of CT scan is decomposed into several individual slices. Before slicing the input volume, the individual CTs are normalized to fit into a common intensity range between -1000HU to 2000HU. For decomposing volume, the slices along the axial direction are considered. Similar to the CT samples, corresponding GT masks are also broken down into slices along the axial direction with a normalized intensity range from 0 to 1.

Network Training

The performance of any SLS network is heavily dependent on the initialization of the network weights. A wrong choice of the initialization fails to handle the highly non-linear data. A robust initialization method assures deep investigations for a model developed from scratch. Here, He initialization is used. He initialization computes the network weights from the Gaussian probability distribution function. The network weights are the function of the SD. SD varies with the number of inputs to the nodes (N), such that $SD = \frac{2\pi}{\sqrt{N}}$. The model weights are updated using standard backpropagation with “Adam” optimizer.



The model has been trained over a set of 19487 2D images from 78 volumetric CT scans. The validation set comprises 2166 images from the LUNA dataset. The proposed model has been trained using two different loss functions viz. dice loss and “binary cross-entropy loss (BCEL)” to investigate the best. The model training is headed towards minimizing the loss incurred due to incorrect predictions. The dice loss function is defined as

$$J(w)_{Dice} = 1 - \frac{2 |f(x; w) \cap y|}{|f(x; w)| + |y|} \quad (3)$$

Where, $f(x; w)$ represents the predicted value as a function of network weights w , and y indicates the GT image. The time required to train the model on a machine with 4GB RAM is 2776 s/epoch. The average time required for sample prediction is 75ms. The BCE loss is defined as

$$J(w)_{BCE} = -\frac{1}{N} \sum_{i=1}^N y_i \log(f(x_i; w)) + (1 - y_i) \log(1 - f(x_i; w)) \quad (4)$$

This model has been trained and evaluated for different combinations of the loss functions and performance indices.

V. RESULTS AND DISCUSSION

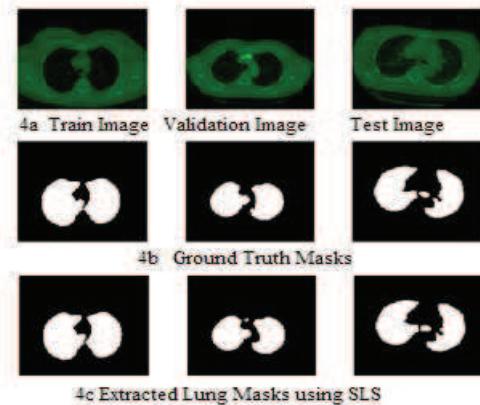


Fig. 4 Lung Segmentation Results using SLS: modified 2D Unet

TABLE 1 SLS MODEL PERFORMANCE FOR DIFFERENT PARAMETERS

	E-I	E-II	E-III
Loss Function	Dice Loss	Dice Loss	BCEL
Metric	Accuracy	DSC	Accuracy
Training Accuracy	0.994	0.9815	0.9967
Validation Accuracy	0.9977	0.9917	0.9966

Table 1 summarizes the SLS model performance for different combinations of loss functions. The performance of the network does not project any significant difference. Table 2 compares the performance of all three techniques on some randomly selected sets of images. The reliability of the T-I and T-II on the OT results in poor segmentation. It is observed from Table 2 that the T-II performs superior as compared to T-I and SLS model outperforms and achieves the highest values of DSC and JI.

TABLE 2 COMPARISON OF DIFFERENT LS TECHNIQUES

Image	DSC			Jaccard Index (JI)		
	T-I	T-II	SLS	T-I	T-II	SLS
1	0.900	0.932	0.990	0.818	0.874	0.980
2	0.962	0.987	0.994	0.928	0.974	0.998
3	0.975	0.994	0.996	0.951	0.988	0.992
4	0.853	0.968	0.991	0.744	0.938	0.983
5	0.959	0.985	0.994	0.921	0.971	0.990
6	0.933	0.972	0.986	0.875	0.945	0.973
7	0.890	0.975	0.988	0.802	0.951	0.977
8	0.968	0.984	0.995	0.939	0.969	0.991
9	0.945	0.979	0.996	0.896	0.960	0.993

Fig. 4 displays the segmentation results using SLS for the sample image from training (column-I), validation (column-II), and test (column-III) sets. 4b and 4c represent the GT masks and ELMs respectively.

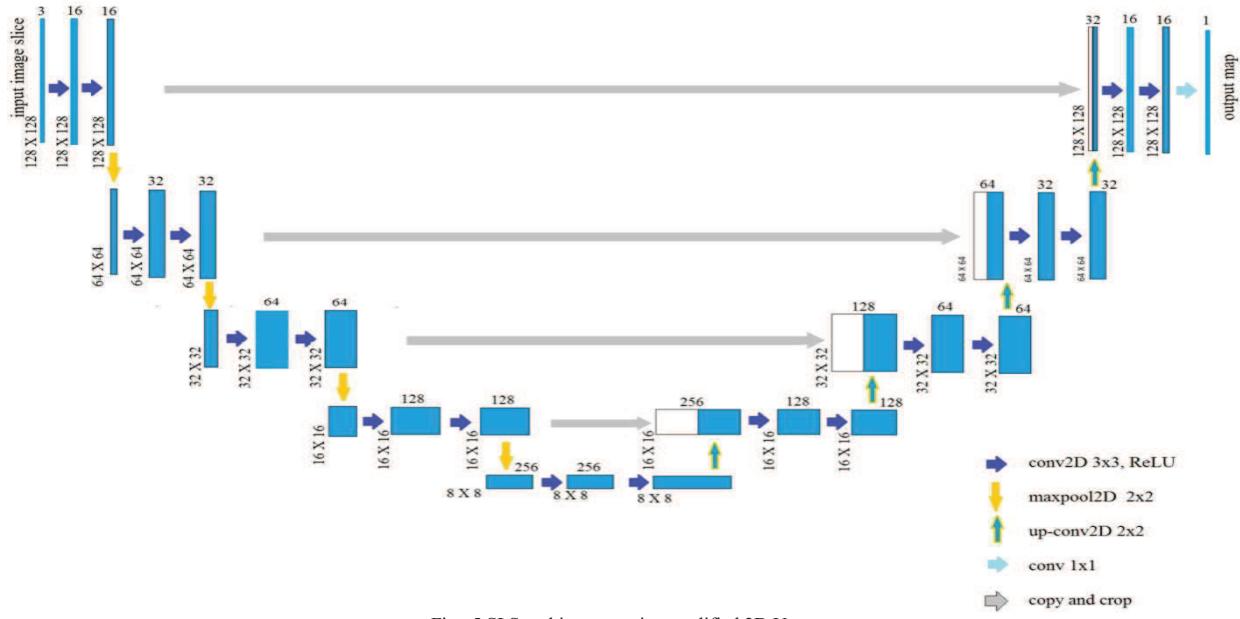


Fig. 5 SLS architecture using modified 2D U net

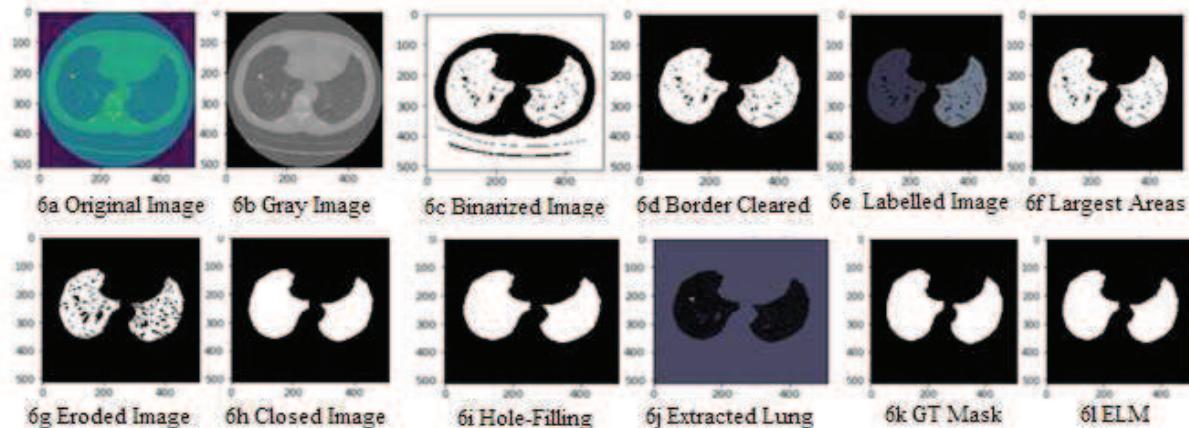


Fig. 6 Segmentation results using T-I

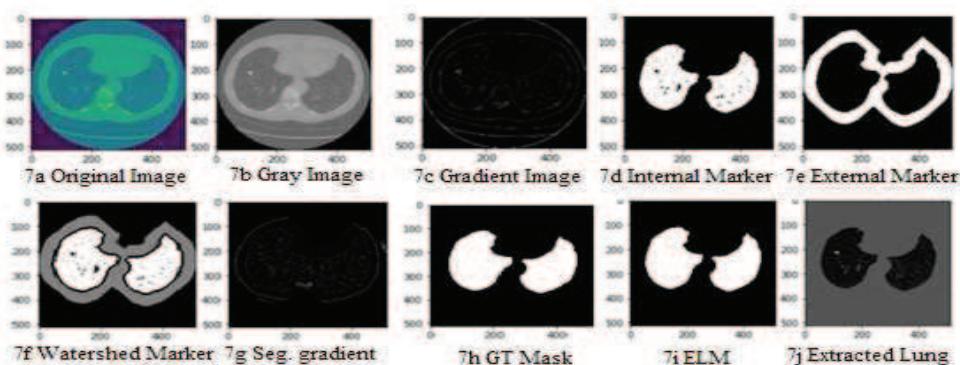


Fig. 7 Segmentation results using T-II

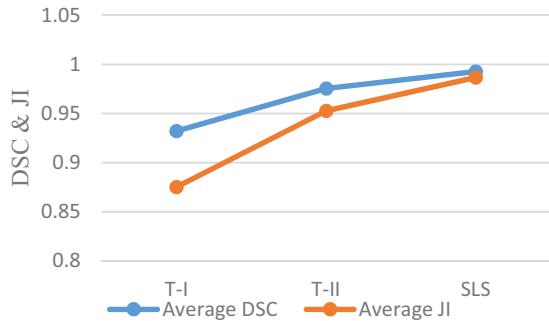


Fig. 8 Comparison of different LS techniques

Fig. 6 and Fig.7 showcase the resultant images while segmenting using T-I and T-II respectively. Each sub-image represents the result of subsequent complex image processing operations performed on the input CT slice. Fig.8 compares the above-mentioned techniques in terms of average DSC and average JI on a common dataset. The SLS architecture has achieved the highest values of DSC (0.9927) and JI (0.9867) respectively as compared to T-I and T-II.

VI. CONCLUSION AND FUTURE WORK

In a volumetric scan of the lung, the area useful for clinical investigations is not more than half a CT. A lot of computational time will be reduced if we automate the process of extracting the lung parenchyma from the whole CT volume. Such segmentation will improve the speed of investigations and consequently reduce the region of search space. This evokes a significant drop in false detections. In this study we tried to evaluate some of the LS approaches on a common database. We implemented a novel method for parenchyma segmentation using a modified 2D U-Net model.

The SLS automated the process of LS and has a sample segmentation time of only 75ms per image. This technique overcomes the trivial issue of OT selection, inevitable in most of the recent LS techniques. This model offers the precise localization of pixels while segmenting. This model can be trained on a low configuration machine with a small amount of data. The SLS model makes the predictions robust against various human interventions such as fatigue, overload, missing steps, etc. SLS outperforms T-I and T-II and achieves the highest segmentation accuracy of 99.77%. This technique would effectively reduce the region of search space and help the clinicians for diagnosing the majority of lung disorders. A 3D segmentation of lung volume can be constructed by stacking all the individual segmented slices together. This approach will be further extended to excavate region-specific pulmonary nodules.

REFERENCES

- [1] R. D. Babu R, R. M. Shankar, and B. N. Rao, "Automated Segmentation of Lung Regions using Morphological Operators in CT scan," International Journal of Scientific & Engineering Research, vol. 4, no. 9, pp. 1114–1118, 2013, [Online]. Available: <http://www.ijser.org>
- [2] M. S. Rahman, P. C. Shill, and Z. Homayra, "A New Method for Lung Nodule Detection using Deep Neural Network for CT Images", 2nd International Conference on Electrical, Computer And Communication Engineering, ECCE 2019, pp. 1-6, 2019, DOI: 10.1109/ECACE.2019.8679439.
- [3] L. Y. Tseng and L. C. Huang, "An adaptive thresholding method for automatic lung segmentation in CT images," IEEE AFRICAN Conference, no. September, pp. 1–5, 2009, DOI: 10.1109/AFRCON.2009.5308100.
- [4] N. S. Nadkarni and S. Borkar, "Detection of lung cancer in CT images using image processing," Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, vol. 2019-April, no. Iconic, pp. 863–866, 2019, DOI: 10.1109/icoei.2019.8862577.
- [5] N. Mesanovic, M. Grgic, H. Huseinagic, M. Males, E. Skejic, and S. Muamer, "Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm," 18th International Conference on Systems, Signals and Image Processing-IWSSIP, pp. 395–400, 2011.
- [6] H. S. Kim, H. Yoon, K. N. Trung, and G. S. Lee, "Automatic Lung Segmentation in CT Images Using Anisotropic Diffusion and Morphology Operation," pp. 557–561, 2008, DOI: 10.1109/cit.2007.143.
- [7] T. Peng et al., "Hybrid Automatic Lung Segmentation on Chest CT Scans," IEEE Access, vol. 8, no. 3, pp. 73293–73306, 2020, DOI: 10.1109/ACCESS.2020.2987925.
- [8] F. Liao, M. Liang, Z. Li, X. Hu, and S. Song, "Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3484–3495, 2019, DOI: 10.1109/TNNLS.2019.2892409.
- [9] L. Zhao, "3D Densely Connected Convolution Neural Networks for Pulmonary Parenchyma Segmentation from CT Images," Journal of Physics: Conference Series, vol. 1631, no. 1, 2020, DOI: 10.1088/1742-6596/1631/1/012049.
- [10] C. Liu, R. Zhao, and M. Pang, "Lung segmentation based on random forest and multi-scale edge detection," IET Image Processing, vol. 13, no. 10, pp. 1745–1754, 2019, DOI: 10.1049/iet-ipr.2019.0130.
- [11] J. Pedrosa et al., "LNDb: A lung nodule database on computed tomography," arXiv, pp. 1–12, 2019.
- [12] S. G. Armato et al., "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans," Medical Physics, vol. 38, no. 2, pp. 915–931, 2011, DOI: 10.1111/j.13528204.
- [13] E. Papavasileiou, M. Prokop, M. Saletta, C. M. Schaefer-prokop, and E. T. Scholten, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge".
- [14] W. Weng and X. Zhu, "INet: Convolutional Networks for Biomedical Image Segmentation," IEEE Access, vol. 9, pp. 16591–16603, 2021, DOI: 10.1109/ACCESS.2021.3053408.

Private Messaging Service using AES Encryption and Toxicity Detection

1st Shanmuga Priya R

*Department of Information Technology,
Madras Institute of Technology
Chennai*

shanmurajendran2@gmail.com

2nd Tharunraj M

*Department of Information Technology,
Madras Institute of Technology
Chennai*

tharunoptimus@outlook.com

3rd Rozen Berg D

*Department of Information Technology,
Madras Institute of Technology
Chennai*

rozenberg4christ@protonmail.com

4th Sanjay Kannan M

*Department of Information Technology,
Madras Institute of Technology
Chennai*

sanjaykannanmurali2001@gmail.com

Abstract—With the advancement of digital technology in recent decades, the style of communication and use of digital accessories in our daily lives has changed dramatically. It is undeniable that the introduction of the mobile phone/smart phone has improved our standard of living and made life easier. A private messenger is an application that enables for encrypted communication. To provide safe data transmission between the sender and recipient, a protected private instant messenger is required. The messenger gives main emphasis on textual value by checking the text toxicity using a ML model in a novel way by not storing metadata of user at any juncture without compromising the quality of messaging service which is not present in main stream messaging application. The architecture of private messengers is presented in this work. The proposed architecture was created and tested on a private messaging application, and the findings show that it can increase data security for messenger applications.

Index Terms—privacy, messenger, encryption, communication

I. INTRODUCTION

Online messaging can be observed as the latest trend in the field of communication. The main advantage of this method of communication is that people can text each other in real-time. Chatting is different from other communication media in the Internet such as blogs, forums, email due to the fact that the user has a feel of spoken conversation, which is not present in other forms of online communication media. Online messaging can be unicast, broadcast, multicast, etc. Nowadays some application offers any cast and the scope of this mode of messaging is beyond the scope of the current paper. In online messaging the message is sent in an instant to the receiver over the Internet or any other network.

Instant Messaging enables effective and efficient communication by allowing for rapid acknowledgement or response. Instant Messaging, on the other hand, isn't always supported by transaction control. In many cases, instant messaging comes with additional features that might help it become even more popular. User can now use web cameras to see others over Internet for free of cost and they can speak each other through microphones and hear the other person through any audio device. Many programmers offer file transfers, albeit the file size that may be transferred is usually limited. Text based chatting application usually stores the chat for future reference. It is now called a message history which can be synonymous to the storage of emails.

As the use of internet-based messaging applications has grown, many Big Tech companies have seen this as an opportunity to harvest data about the users of that particular chat service. This poses serious privacy problems, as any conversation between two individuals is not protected in any way.

The first requirement for any private messenger is end-to-end encryption, often known as E2EE. This is a non-negotiable prerequisite for any programme claiming to be a private messenger. End-to-end encryption ensures that our messages are encrypted along their journey, ensuring that even if they are intercepted, no one will be able to read them. The encryption key required to read any given message is held by just us and the person to whom we are messaging. We must also evaluate who is in charge of receiving and preserving our messages. The firm that runs the messaging software may be able to decipher our messages if it stores them. Law enforcement agencies or other third parties may be able to force a corporation to divulge our messages based on where and how they are stored. To overcome this, several messengers avoid keeping messages in one location or set messages to expire after a certain period of time.

Furthermore, while some messengers are private, they are not anonymous. The contents of a message we send will usually be hidden by a private messenger, but the fact that we sent it will not be hidden. Not only does an anonymous messenger hide the contents of our messages, but it also hides all of the information about the interaction. The proposed project constructs architecture for anonymous chat application using web sockets. The web sockets are constructed using connection oriented architecture. The protocol used for the connection oriented is Transmission Control Protocol (TCP). The data sent is secured using E2EE encryption which has been a globally accepted encryption standard used by messaging platform like WhatsApp.

II. RELATED WORK

There are many researches carried out regarding anonymous and privacy related chat messengers. Dijana Vukovic et al. [1] argue how Internet communication privacy has been harmed. CryptoCloak, a programme for chat communication privacy protection, was created by them. For chat communication, CryptoCloak provides privacy protection. Encrypted communication is hidden behind a lively, low-cost chat interaction. Mass surveillance

spying engines are uninterested in communication conducted in this manner. According to their research, we are living in the era of Internet Surveillance.

Botha et al. [2] have distinguished between pure secure messaging apps and messaging apps which have a poor level of security. Their report provides us the evaluation of the applications based on privacy features and security levels. The report also gives us a insight that most of the applications stores the personal data and metadata of the users, while consumers desire protection against surveillance as well as hacking attacks. It was also mentioned that in March 2016 a Wikileaks revelation mentioned that CIA has broken into Whatsapp and Signal's security doors. [3] According to a Guardian story from 2017, Facebook, the company that owns Whatsapp, got the access to all the encrypted messages due to a critical security flaw. [4]

Endeley et al. have outlined the benefits of keeping E2EE as the default encryption system in Whatsapp and other messaging applications .It also stressed that no agency should have the access to the conversation of the users. Their study also discusses the advantages of encryption. They claim that E2EE makes data between communication parties safe, eavesdropper-proof, and difficult to decrypt. The technology gives the user a safe haven from third parties, hackers and other bodies who try to steal their data when it is in transit between various nodes over the network. [5]

Telegram's security protocol has been studied by Jeeun Lee et al. Telegram employs a unique protocol called MTProto, according to them. In cloud conversations, it offers a client and server encryption which is synced between all the devices used by the user, E2EE is used for secret chats feature in the messaging application. [6] The authors focused on Advanced Encryption random padding vulnerabilities and known exploits, padding length extension, and last block substitution. Since these flaws were theoretically demonstrated in 2015, they attempted to mimic attacks by encrypting conversations between Alice and Bob. Building blocks for E2EE in general IM systems are required so that further attacks or cryptographic strength measures may be handled easily while retaining the basic design of cryptographic protocols is maintained. For further testing of random padding vulnerabilities, a simplified MTProto was created as a preliminary step. New cryptographic protocols are difficult to design, and ensuring their reliability takes a long period. MTProto was created in a secure manner, according to Telegram, despite some cryptographic primitive flaws.

Following previous data breaches, online consumers have been concerned that their private information on the Internet is vulnerable to misuse, according to Steven Song et al. [7] These concerns prompted data privacy rules such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), which aim to protect personal data from unauthorized access. In a related announcement, Facebook said it would adopt end-to-end encryption (E2EE) across its entire social media messaging services. E2EE restricts communicating parties' access to private messages. These encrypted discussions are not accessible even to the firm that manages the messaging system. They claim that the implementation of E2EE, on the otherhand, has its own set of issues. E2EE, for example, has the potential to promote terrorist operations because it protects not just the privacy of innocent users but also the privacy of terrorists. The software speed was monitored when the usage of AES-

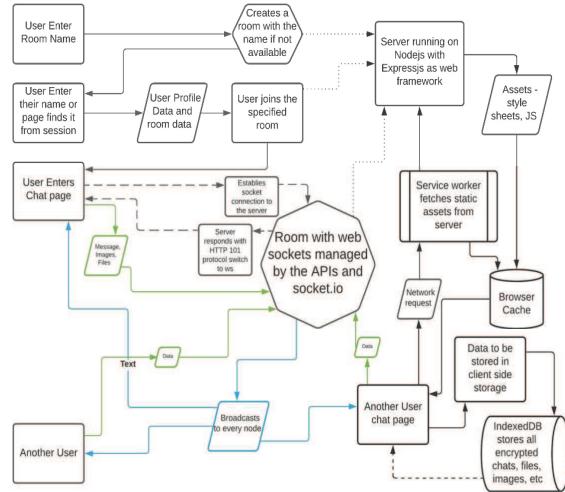


Fig. 1. Proposed Architecture.

128 encryption is invoked at the ends of the user and receiver and the results were presented in a work by Dag Arne Osvik et al. [8] They have used low class hardware devices at one end which includes 8-bit AVR micro controllers and 32-bit ARM microprocessors and high class hardware devices such as NVIDIA GPUs at other end .They obtained the results using platform dependent methods and the process is explained in depth in their report.

Noura Aleisa compared three encryption standards: 3DES, AES, and DES. Although it has been said that DES is insecure and no longer useful, this is still not practical because these algorithms cannot be broken easily by any modern computers. However, the author claims that as computer power grows, efficient algorithms are needed to secure the system from hacker attacks. AES is an encryption system that works very swiftly in most of the devices

. According to the report, AES will provide higher security in the long run due to its larger block size and longer keys. [9]

The following architecture in Figure 1 is used in the chat application. The components used in the architecture are explained in the following modules. The data flow is shown in the architecture diagram when a user tries to send a message in the chat room.

III. PROPOSED WORK

The following subsections explain the components used in the chat application. The next section will explain how the components are used.

A. Express JS

It is a fast, unopinionated, minimalistic web framework for node.js. With the package manager npm, ExpressJS makes it easy to use the building blocks to develop our server. It uses middleware to simplify the process of data transfer between the server and the client. It takes care of the API endpoints out of the box allowing the developers to focus on the logic of the server side requests handling. It made node.js development a breeze and added a slew of new features. Every API requests fire an event, for the endpoint, the ExpressJS responds to it asynchronously to make sure it does not block the executing thread. It has a very

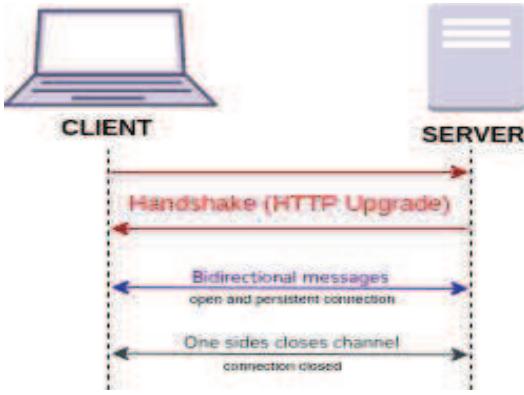


Fig. 2. Connection using WebSocket.

easy implementation of web sockets which is used in this Privacy Talk.

B. WebSocket

It is a telecommunications networking protocol that facilitates low latency communication between a server and a client. According to figure 2: WebSocket, unlike HTTP, allows for full duplex communication. WebSocket also allows us to send messages in streams over TCP. TCP is the only protocol that deals with the streams of data and has no idea of a message. Due to the TCP handshakes and overhead of HTTP headers, the methods like DHTML, Cross Frame Communication, HTTP Polling, Live Connect, Forever Frame, AJAX, HTTP Long Polling, XHR Streaming were expensive and difficult to implement for the Privacy Talk.

The Websockets take up a different approach by leaving a connection open to allow bi directional data transfer. The Client will send a HTTP request to the server asking to open a WebSocket connection. If the server agrees, it will send a 101 Switching Protocols response, at which point, the handshake is complete. The TCP IP connection is left open, allowing bi directional messages to pass between the two parties with very low latency. The connection will stay open until one of the parties drops off, then the TCP resources can be unallocated. It's often referred to as a full duplex connection, which is a telecommunications term that defines how a phone line works where both parties can send messages or talk at the same time.

C. Socket.io

It is the websockets implementation for node.js allowing developers to quickly start working with the protocol. It is a library that allows a server and a local web browser to establish a connection in both directions and at any time. It consists of the following components:

- a server in Node.js
- The library that is loaded in the browser either using script tags or a bundler like webpack

Privacy Talk implements the sockets events with the library at the server level making the browser to server communication possible. It uses the websockets, falling back to HTTP long polling strategy thereby increasing its reliability in a real world environment. The library has the advantages over the node.js implementation of websockets making the development and debugging of Privacy Talk easier:

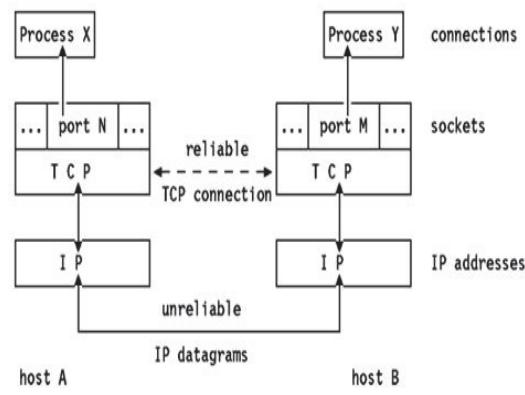


Fig. 3. TCP Socket Connection.

- reliability
- acknowledgements
- multiplexing
- packet buffering
- reconnection by itself
- broadcasting to all customers or a selection of customers

D. Transmission Control Protocol

Transmission Control Protocol is considered as the most important protocol in the field of computer networks. TCP ensures that a stream of octets is delivered in a reliable, orderly, and errors are made sure that they are checked between the sender and receiver who are connected using the IP addresses. TCP has been used for various internet based applications such as email, file transfer , etc.

TCP is connection-oriented, which gives us the view that data can be transferred only after the establishment of a connection between client and server. The server always listens for any incoming connection requests from any client. The three way handshake mechanism and error detection methods make us depend on the protocol for any uses. Figure 3 illustrates about TCP Socket Connection. The figure shows us that we need IP datagram to have a socket connection. The port number along with the IP address will help us establish a full TCP Socket Connection. The figure also helps us visualize how different processes in different hosts are connected with each other with the help of reliable data transfer methodology.

E. AES encryption

AES is a Rijndael block cipher variant developed by two Belgian cryptographers. [10] Rijndael is a cipher family with different key and block sizes. The authors had submitted their proposal to NIST for the AES key selection. NIST chose members which had a block size of 128 bits with different key lengths such as 128, 192, 256 bits. It replaced data encryption standard (DES).AES is a symmetric key algorithm which implies to us that similar keys are used to encrypt the data and also decrypt it.

Advanced Encryption Standard (AES) encryption as shown in Figure 4, often known as 256-bit AES encryption, is a data/file encryption security approach that encrypts and decrypts data or files using a 256-bit key. It is one of the most trustworthy encryption algorithms. Figure 4 also helps us visualize how the encrypted message even in the hands of harmful people cannot



Fig. 4. AES Encryption.

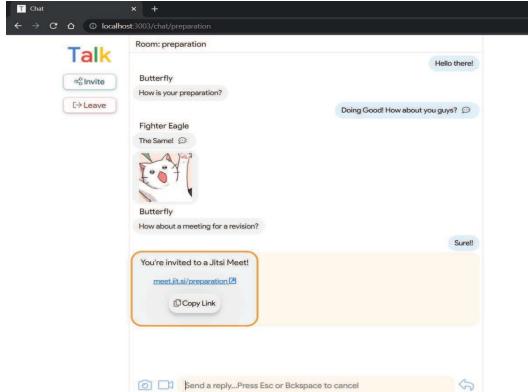


Fig. 5. Chat application in usage.

cause any harm to the owner of the message since the encrypted message's cypher cannot be broken by any third party member other than the user and sender.

F. IndexedDB

IndexedDB is a NoSQL storage solution with a vast size. It allows you to save almost anything in the browser of the user. IndexedDB provides transactions in addition to the standard search, get, and put operations. IndexedDB is a client side API for storage of huge bytes of data which should be structured like blobs, etc. This API makes use of indexes to allow for fast data searches. DOM storage performs poorly when compared to IndexedDB for storing huge bytes of structured data.

Each IndexedDB database belongs to a single origin, which means it can't access or be accessed by anybody else. IndexedDB is similar to a SQL database management system and also a JavaScript based object oriented database. IndexedDB stores and retrieves object based on a key, any object that can be stored using the structured clone algorithm is supported. We must first specify the database schema, then connect to it and obtain and update data through a series of transactions.

G. Text Toxicity Checking

This is a pre-trained machine learning model based on Universal Sentence Encoder. This is invoked using tensorflow.js. The model predicts that if the text entered by the user contains any

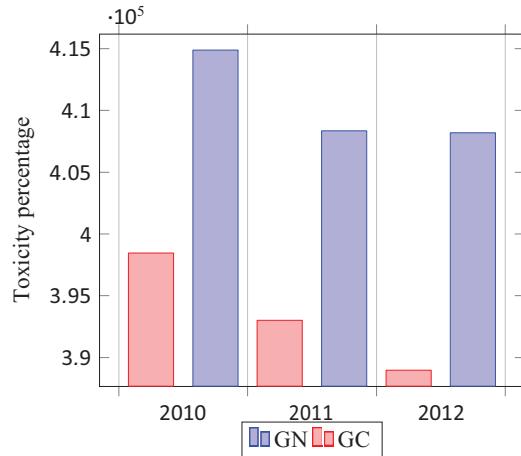


Fig. 6. Graph showing the significance of toxicity percentage

```

function getMessageEncoding() {
    const messageBox = document.querySelector(".aes-ctr #message");
    let message = messageBox.value;
    let enc = new TextEncoder();
    return enc.encode(message);
}

function encryptMessage(key) {
    let encoded = getMessageEncoding();
    // counter will be needed for decryption
    counter = window.crypto.getRandomValues(new Uint8Array(16));
    return window.crypto.subtle.encrypt(
        {
            name: "AES-CTR",
            counter,
            length: 64
        },
        key,
        encoded
    );
}

```

Fig. 7. Data Encryption using AES.

toxic content such as threatening language , insults , hate words, etc. This model will help the children and younger people as it can protect them from harmful internet users. This also helps the chat application to maintain online safety. The figure above shows the toxicity percentage of GN (annotators with no context) and GC (annotators with context) taken in three years.

IV. FUNCTIONING OF THE COMPONENTS

Privacy Talk is a private messenger which uses web sockets which facilitates full duplex communication using a single TCP Connection. The web application runs a server on Node.js and Express provides the web framework required for the project. The nodes connected with the web application uses RESTful APIs for every conventional request over HTTP protocol. For seamless UI and satisfactory user experience the web application will be enhanced using jQuery, JavaScript, CSS and PUG for frontend templating. The project will also use DNS protocol to find the required IP address, FTP protocol for file transfer and HTTP/2 for the transfer of static assets. The expected browser technologies to be used are Service Workers for periodic background syncs, Indexed DB API for managing the JSON objects in a NOSQL database within browser and Background sync API that lets the application defer action until the user has stable

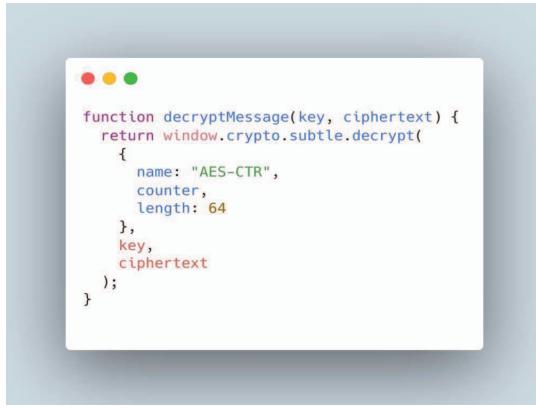


Fig. 8. Data Decryption using AES.

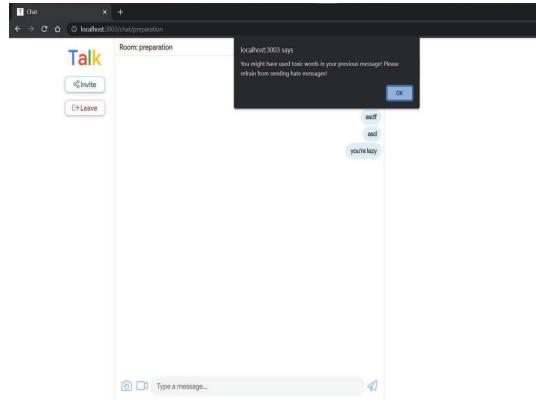


Fig. 10. Text Toxicity Prediction Model.

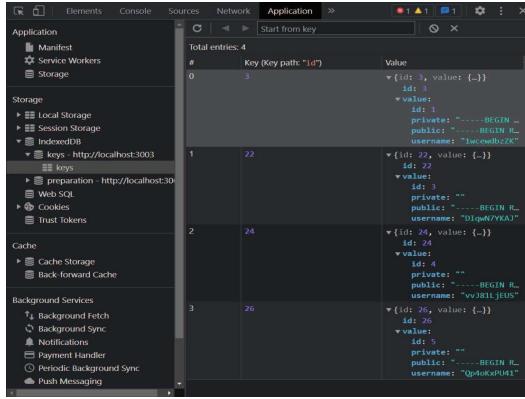


Fig. 9. IndexedDB storing the keys .

connectivity. The web socket with TCP act as the medium to share text, links, files and other blob content.

V. RESULTS

A. Chat Component Output

The Figure 5 shows the chat application being used by different users to chat privately. We can confer from the figure that we can send text messages to the whole chat room. Sending of images is also feasible with the proposed chat application using the defined architecture. The users are also given a chance to start up a video call using open source software, Jitsi Meet.

B. AES Encryption Code Snippet

The code snippet which encrypts the messages sent is shown in Figure 6. We can see how the chat application has implemented AES-CTR encryption using JavaScript. There is also a counter variable to prevent any thrashing due to encryption. The code snippet which encrypts the messages sent is shown in Figure 7. It is clear that we need to give the same variables that we gave in Figure 6. The reason behind this procedure is due to the usage of AES-CTR encryption. The length value should also be the same value given in the encryption phase.

C. IndexedDB Output

The IndexedDB stores the private key and the public key on the browser locally to maintain privacy and the keys are not sent

to any other server or database outside the web browser. It also stores the keys broadcasted to it. It is show in in Figure 8. The storage of encryption keys in local storage using web browser is a core feature of this chat application. This prevents any data from being sent to any server which may pose the risk of breaking the encryption and make the messages' security to be flawed.

D. Text Toxicity Prediction Model Output

The user's input has been checked for text toxicity and the results are shown in Figure 9. This prevents abusers and harmful people from entering the chat room. Maximum usage of toxic words leads to expulsion from the chat room. The figure also shows how this feature can help children remain safe in Internet chat rooms and can be prevented from exploitation by online abusers.

VI. CONCLUSION

In the era of internet communication especially using online messaging services has been under the surveillance of many Big Tech companies and Government Agencies. The inclusion of privacy is now a very big factor to the modern users of chat applications. The proposed architecture in the paper includes a anonymous oriented way of speaking with other people over Internet. We can observe that the chat application performs very well and is faster than many conventional chat applications used today. The AES Encryption makes the chat safe and protects intruders from spying an Internet user. The text toxicity model helps the chat model be protected from online abusers and young users away from harmful people and makes the virtual world a safe place for people of every age.

REFERENCES

- [1] D. Vukovic, D. Gligoroski and Z. Djuric, "On privacy protection in the Internet surveillance era," 2014 11th International Conference on Security and Cryptography (SECRYPT), 2014, pp. 1-6.
- [2] Botha, Johnny, W. C. Vant, and Louise Leenen. "A comparison of chat applications in terms of security and privacy." In Proc. 18th Eur. Conf. Cyber Warfare Secur., p. 55, 2019.
- [3] <https://wikileaks.org/ciav7p1/>
- [4] <https://www.theguardian.com/technology/2017/jan/13/whatsapp-design-feature-encrypted-messages>
- [5] Endely, Robert E. "End-to-end encryption in messaging services and national security—case of WhatsApp messenger." Journal of Information Security 9, no. 01 (2018): 95.

- [6] Lee, Jeeun, Rakyong Choi, Sungsook Kim, and Kwangjo Kim. "Security analysis of end-to-end encryption in Telegram." In Simposio en Criptografía y Seguridad Informática, Naha, Japón. Disponible en <https://bit.ly/36aX3TK>. 2017.
- [7] Song, Steven. "Keeping Private Messages Private: End-to-End Encryption on Social Media." BC INTELL. PROP. TECH. F (2020): 1-11.
- [8] Osvik, Dag Arne, Joppe W. Bos, Deian Stefan, and David Canright. "Fast software AES encryption." In International Workshop on Fast Software Encryption, pp. 75-93. Springer, Berlin, Heidelberg, 2010.
- [9] Aleisa, Noura. "A Comparison of the 3DES and AES Encryption Standards." International Journal of Security and Its Applications 9, no. 7 (2015): 241-246.
- [10] Daemen, Joan, and Vincent Rijmen. "AES proposal: Rijndael." (1999).

Application of ANN, SVM and KNN in the Prediction of Diabetes Mellitus

1st M Revathi
Information Technology
St.Joseph's Institute of
Technology
Chennai, India
gmrevathigopinath@gmail.com

2nd A.Beenam Godbin
Research Scholar
Vellore Institute of
Technology
Chennai, India
beenu0424@gmail.com

3rd S.Nikkath Bushra
Information Technology
St.Joseph's Institute of
Technology
Chennai, India
ferozbushra@gmail.com

4th S. Anslam Sibi
Information Technology
St.Joseph's Institute of
Technology
Chennai, India
anslam.sibi@gmail.com

Abstract—Diabetes Mellitus is a serious health problem which affects huge population in world. Age, food habits, improper eating and sleeping schedule, genetic and lack of exercise are major causes for diabetes. Diabetes leads to other major health problems like cardio vascular disease, kidney disease, lung disease etc., If diabetes is not treated properly, it could result in the failure of body organs. Proper diagnosis and treatment of diabetes is important to maintain the health of the patient. Machine learning application in healthcare sector is increasing day by day. Machine Learning helps in the extraction of useful information from raw medical data which is useful for identification of disease. In this paper, Artificial Neural Network (ANN) with varying batch size and epochs and algorithms like K-Nearest Neighbor and Support Vector Machine (SVM) are used in the prediction of Diabetes. ANN shows better performance than other two algorithms.

Keywords—*Diabetes Mellitus, Support Vector Machine, Artificial Neural Network, K-Nearest Neighbor, Classification.*

I. INTRODUCTION

Body's ability to process blood glucose when impaired it is called as diabetes. The deficiency in insulin secretion, insulin action, or both can cause hyperglycemia. Complete depletion of insulin secretion (T1D) cause Type 1 diabetes. when a patient's body is unable to utilize the insulin generated which is Type 2 diabetes (T2D), diabetes spreads so fast. Both types of diabetes are on the rise, although T2D has a greater rate of growth than T1D. 90~95% of all diabetes cases falls under T2D. Even though the exact aetiology of diabetes is unknown, scientists believe that hereditary factors as well as environmental lifestyle have a significant effect. Despite the fact that it is incurable, it can be controlled with treatment and medicine. Diabetes patients can be prone to other health issues like heart disease and nerve damage. Complications can be avoided with early identification and treatment.

Many bio-informatics research fellows made an attempt to address this issue by developing systems and methods that will help in diabetes prediction. Research was conducted in numerous different fields including Machine Learning (ML) and Artificial Intelligence (AI), to lessen the impacts of diabetes and increase the quality of patient care. ML based diabetes prediction algorithms is described in many researches. Identifying current condition (through screening and diagnosis) and forward prediction approaches are two sorts of these strategies. Forward prediction methods use current and previous medical records for forecasting the incidence of diabetes in the future. Current condition identification deal with classification of current data instances. Forward prediction methods use current and

past medical records to forecast the incidence of diabetes in the future.

They either used classification or association algorithms to build prediction models. Most popular algorithms were listed below,

- ✓ Decision Trees
- ✓ Support Vector Machines (SVM)
- ✓ Linear Regression.

Another form of Machine Learning technology is Artificial Neural Networks (ANN) which is well-known for its speed and precision.

Machine learning approaches are commonly employed in diabetes prediction as the outcome is good. Decision trees are a prominent machine learning method which have the capability of good classification. Random forest generates good amount of decision trees. New machine learning technology Neural networks outperforms traditional methods in a various ways.

Method that combines a range of ML algorithms, techniques on data mining and statistical methods for predicting future using present and past healthcare data is called as Predictive Analysis.

ML and regression techniques can be used to perform predictive analytics which strives to improve clinical outcomes by improving diagnosis accuracy, improving patient care, and maximizing resources.

Artificial intelligence elements can be used for computer systems creation which have the capability to learn from prior experiences and there will be no need to program for each and every scenarios.

Machine learning is considered as a critical requirement in today's medical field for reducing manual effort enabling automation. Lab tests conducted using blood glucose under fasting and oral glucose tolerance are currently used for detecting diabetes which is time consuming process.

II. LITERATURE REVIEW

Qawqzeh et al. [1] suggested a model for diabetes categorization which is based on photoplethysmogram analysis called logistic regression. Their proposed approach properly identified 552 people as non-diabetics with a 92 percent accuracy rate. The proposed technique, on the other hand, is not compared to state-of-the-art techniques. Pethunachiyar [2] introduced a machine learning-based diabetic mellitus classification method. Support vector machine with a variety of kernel functions is used in this. He discovered that SVM with a linear function outperformed naive Bayes, decision trees, and neural

networks. For diabetes categorization, Gupta et al. [3] used support vector machine methods and naive Bayes. They employed a feature extraction method and k-fold cross-validation to increase the model's accuracy

A comparative analysis of diabetes classification approaches was published by Choubey et al. [4]. To categorize patients as diabetic or not, they employed AdaBoost, K-nearest neighbour regression, and the radial basis function on both datasets. They also employed LDA and PCA for feature engineering, and it was discovered that both are effective with classification algorithms for enhancing accuracy and deleting undesired features. To categorize and forecast diabetes, Maniruzzaman et al. [5] used machine learning approach. For diabetes categorization four machine learning algorithms are used which are AdaBoost, Naive Bayes, Random forest and Decision Tree.

On the PIMA dataset for diabetic categorization, Ahuja et al. [6] performing analysis by comparision of various machine learning algorithms, such as NB, DT, and MLP.

Fine-tuning and efficient feature engineering, according to the authors, can improve MLP's performance. Mohapatra et al. [7] employed MLP to classify diabetes and reached 77.5% accuracy level on the dataset of PIMA, but they were unable to undertake certain state of comparisons.

Singh et al. [8] proposed a technique called stacking-based ensemble for forecasting type 2 diabetes mellitus.

Employing an ensemble of three commonly used supervised machine learning methods, Kumari et al. [9] proposed diabetes prediction system based on soft computing. For early prediction of diabetes on or before onset stage proposal was given by Islam et al. [10] using data mining techniques such as naive Bayes algorithm, logistic regression and random forest.

Malik et al. [11] compared machine learning strategies and data association in the prediction of early and onset diabetes mellitus in women. Hussain and Naaz [12] conducted a comprehensive assessment of machine learning models for diabetes prediction that were presented between 2010 and 2019. They compared classic supervised ML models against Neural network-based techniques in terms of accuracy and efficiency.

Choudhury and Gupta [13] utilized several algorithms to divide people into two groups: those at high risk and those at low risk. They employed an SVM to create a categorization hyperplane, a KNN classification strategy for grouping fresh data, the LR binary classifier method and the DTs, RF, NB classifiers.

Dalakleidi et al. [14] employed 2 set of data: PID, scenario 1 and Hippokrateion, scenario 2. The PID is divided into equal percent of training and testing. Hippokrateion has 70 percent training and 30 percent testing split.

They used logistic model tree algorithm (LMT) and binary logistic regression (BLM), which is a combined simple model version of DT and LR learning.

Harris et al. [15] did medical diagnosis using weighted linear regression for identification of diabetes mellitus (NIDDM) which is non dependent on insulin.

R language was used by Ameena and Ashadevi [16] to create a model based on DTs, SVM, LR and RF where a

sample of 768 women all over the age group of 50 was chosen.

On two datasets, Daanouni et al. [17] proposed a method that uses KNN with algorithm of DT, the first with 2000 instances and the second with 768. The model was trained using eight features or attributes, including BMI, glucose, sugar content in blood and pregnancy. The remaining 20% was utilized to test and the balance 80% was to train. SVM, NB, and DT classifiers were employed by Sisodia D and Sisodia DS [18]. The classification is done using the PIMA Indian diabetes dataset, which is based on UCI's PIMA Indian diabetes dataset.

Xu and Wang (2019) proposed a type-II DM prediction model on probability based on the Learning Process using more than one classifier. XGBoost classifier is used to classify that can be validated with PIDD DM set of data and the forest-based weighted feature which is a random selection technique (RF-WFS) is utilised for the better features gathering. The model's accuracy, specificity, and sensitivity are respectively 93.75 percent, 94.8 percent, and 91.79 percent.

Prabhu et al. (2019) [21,22,23] proposed method for detecting and diagnosing adaptive DR based on presence of retinal nodules, which is a DR symptom. Optic disc is isolated from part of hollow organ picture for several reasons. First, its brightness is similar to that of the brilliant lesions. The qualities of exudates are obtained by extracting them.

III. METHODOLOGY FOR CLASSIFICATION

Figure 1 shows the basic steps involved in classification by Supervised Machine Learning algorithms.

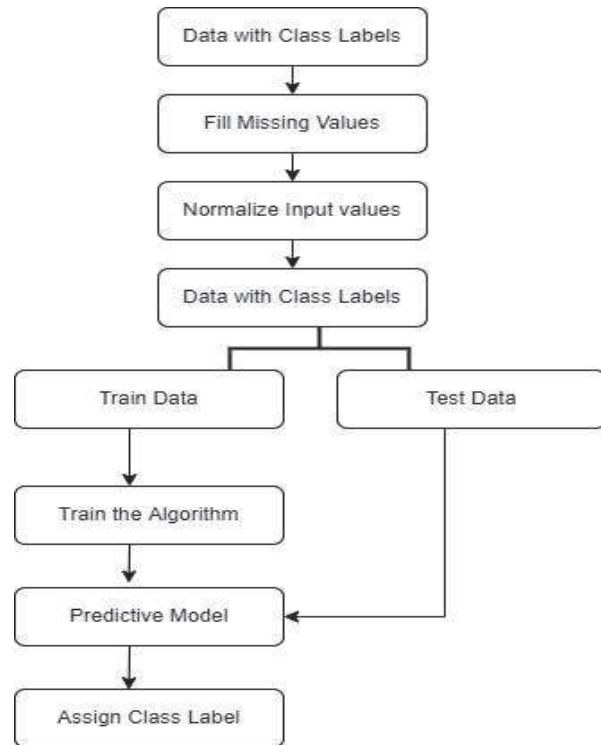


Fig. 1. Steps in Classification

Supervised Machine Learning: This type of ML algorithms takes data with labels as input i.e. the data whose class / category is already known. The machine will be trained first using this labelled data so that the machine will be able to predict the class of a new instance whose label is unknown. Classification and Regression are the types of Supervised ML algorithms.

Classification is a technique which takes the new instance as input and it is able to assign the new instance to a particular category. Example of classification are Support Vector Machine, K-Nearest Neighbor classification and Decision Trees. Regression is a type of Supervised ML algorithms which tries to find the relationship between the input variable and the output variable. Using the relationship, regression algorithms finds the output value for a new input value. Example of regression are Linear regression and Logistic regression.

Unsupervised Machine Learning algorithms takes data without label and analyze those data. This algorithm aims at finding the hidden patterns inside the data. Clustering is an example for unsupervised machine learning algorithms. Clustering takes data with unknown labels as input and tries to find the similarity between data points. K-means clustering and Hierarchical clustering are some examples of clustering.

A. Support Vector Machine

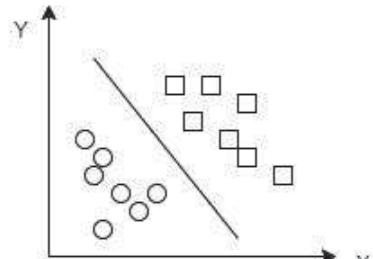


Fig. 2.Hyper Plane in a 2-Dimensional Space

The machine learning algorithm that takes labeled data as input and produces better accuracy in classification as well as regression problems is Support Vector Machines (SVM) [24]. If dataset consists of n attributes (features), each of the data instances is plotted as a point in a plane with n dimensions or n axis . Many lines can be drawn in the plane with n dimensions. The main goal of the SVM classifier is to find a best line that distinguishes or separates one group of similar objects from another group of similar objects as shown in figure 2.

A set of mathematical functions called Kernels[25] is used in SVM for transforming input data into required form. There are different types of Kernels in SVM.

i.Linear Kernel: It is used when the input dataset has many features and the instances are linearly separable. It is

calculated as dot product between any two instances and it is given by equation 1,

$$K(x, x_i) = x * x_i \quad (1)$$

ii.Polynomial Kernel: It is the more generalized form of linear type and it is used when input data set is of non-linear type. The function is given by equation 2,

$$K(x, x_i) = 1 + sum(x * x_i)^d \quad (2)$$

iii.Radial Basis Function Kernel: It is the most commonly used kernel in SVM and it is used for non-linear data and it best separates the classes of objects even when there is no prior knowledge of data. The value of gamma can be from 0 to 1 and the function is given by equation 3,

$$K(x, x_i) = exp(-gamma * sum(x - x_i^2)) \quad (3)$$

B. Artificial Neural Network

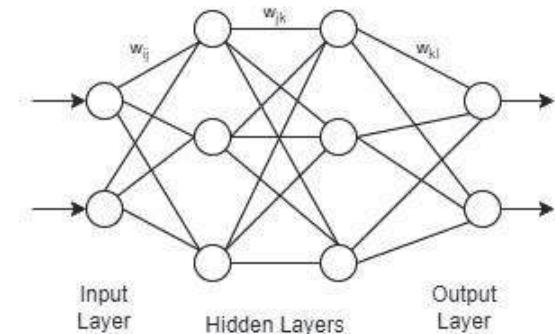


Fig. 3. A Simple ANN with 2 Hidden Layers

Artificial Neural Network(ANN) [24] is developed as a resemblance of the structural and functional organization of the brain in human being. The cell nucleus which is found in count of billions in human brain is referred as neurons and which are connected to each other using axons. Synapses are present in brain which is used to transfer signals among neurons and dendrites are used to receive input.

ANN is composed of many neurons which are organized into layers. The neurons in brain corresponds to nodes/neurons in neural network, synapses corresponds to weights, dendrites corresponds to input and axons corresponds to output. There are three types of layers in ANN: first layer is input layer (only one), next is hidden layer (one or more) and last layer is output layer (only one). A simple ANN with 2 hidden layers is shown in figure 3.

i) Input layer: The input layer is responsible for collecting raw information from the environment which can be of any type such as text, image, signal etc., It passes the value to the hidden layers along with an associated weight.

ii) Hidden layer: The hidden layer performs some complex mathematical calculations on the input data by applying transfer function and activation function. It is used to extract the patterns that are hidden in the input data.

iii) Output layer: The output layer gives the resultant classification based upon the values that are received from the hidden layer.

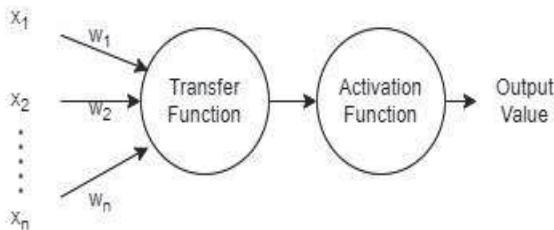


Fig. 4. Working of a Single Neuron

All the input values along with the associated weight is given as input to each neuron. The neuron first applies transfer function on the input data using the formula in equation 4,

$$f(X) = \sum_{i=0}^n (x_i * w_i) + bias \quad (4)$$

The activation function like ReLU, Sigmoid, Tanh etc., is then applied to the output of the transfer function. Based on the output of the activation function, the neuron will decide to fire or not. If neurons are fired, the output will be sent to next layer. The following are some types of activation functions:

i. Binary step function: Here, the output from the transfer function will be compared to a threshold. If it is greater than or equal the threshold, the neurons will fire. The function is given by the equation 5,

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (5)$$

ii. Sigmoid function: It takes any range of values as input and maps it to a value between 0 and 1. This type of function is used when we need the output as a probability value as equation 6.

$$f(x) = \frac{1}{1+e^{-x}} \quad (6)$$

iii. Tanh function: It takes any range of values as input and maps it to a value between -1 and 1 as given in equation 7.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (7)$$

iv. ReLu function: It uses the concept of derivation and it can also be used for back propagation. It produces more efficient outputs when compared to other functions using equation 8.

$$f(x) = \max(0, x) \quad (8)$$

C. K-Nearest Neighbor (K-NN) Classification

The supervised machine learning algorithm which shows good performance is K-Nearest Neighbor Classification [25] and it relies on the knowledge that same group of objects (similar) will appear close to each other. The similar objects will have smaller distances between them and dissimilar objects will have larger distances between them. To find the distance between objects, mathematical formulas like Euclidean distance is used. If (x_1, y_1) and (x_2, y_2) are two instances, the Euclidean distance is calculated as given in the equation 9,

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (9)$$

In K-NN, if an unknown instance comes into the system, the distance between unknown instance and all the existing instances will be calculated. The existing instances will be arranged in the ascending order of distance between the new instance and existing instances. The top K existing instances will be chosen and the majoring class label among the K instances should be identified. The identified majority class label is then given as class label to the unknown instance.

The other name for K-NN classifier is lazy learner, since no prior training will be done and the process starts only after the new instances arrives.

IV. RESULTS AND DISCUSSION

A. Dataset

The input dataset form diabetes prediction is taken from UCI Machine Learning repository. The repository consist of 9 attributes and 768 instances. Out of 768 instances, 500 instances are without diabetes and 268 instances are with diabetes. Attributes such as glucose, pregnancies, skin thickness, BP (pressure level in blood), BMI (value of weight in ratio to height), insulin, Age and likelihood of diabetes (diabetes pedigree function) are used to calculate the prevalence of diabetes.

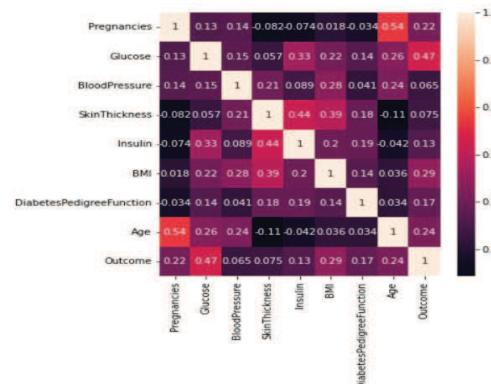


Fig. 5. Correlation Values

Figure 5 shows the dependency value that exists between attributes used in the dataset. From the correlation values (dependency values), we can find the attributes which is more helpful in predicting the output.

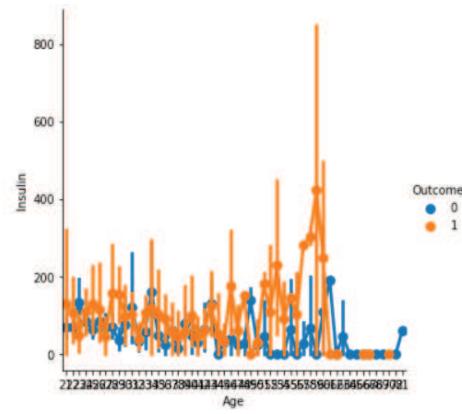


Fig. 6. Age and Insulin Level of Patients

Figure 6 shows the age and insulin level of people with and without diabetes.

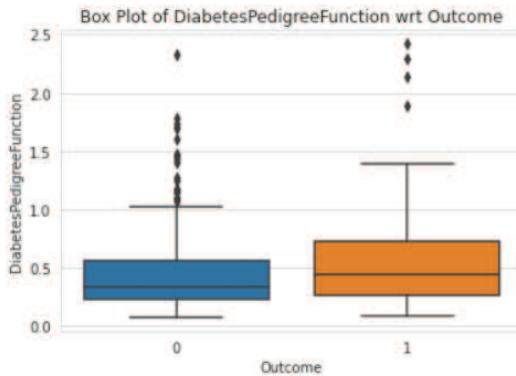


Fig. 7.Impact of Diabetes Pedigree Function

Figure 7 shows the boxplot of the value of diabetes pedigree function for the patients with CKD and patients without CKD. From the figure, it can be observed that the value of diabetes pedigree function for patients without CKD is below 0.5 and the value of diabetes pedigree function for patients CKD is above 0.5.

B. Parameters for Evaluation

i)Accuracy

The performance of classifiers vary depending on numerous number of parameters. So, we need a mechanism to find the efficiency of the classification model. One such useful metric is accuracy. Accuracy can be calculated with dividing the value of correct prediction by the value of total predictions as given in equation 10.

$$\text{Accuracy} = \frac{\text{Number of Correct Prediction}}{\text{Total Number of Input Samples}} \quad (10)$$

ii .Precision, Recall and F1-Score

Precision is one of the important metric and is found by dividing the number of actual true instances by the total number of instances it predicts as positive as given in equation 11.

$$\text{precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (11)$$

Recall is the value of positive instances that are correctly labeled divided by the total instances that are labeled as positive as shown in equation 12.

$$\text{recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (12)$$

F1-Score is defined as weighted mean that is calculated based on the value of precision and recall and is given by the equation 13.

$$\text{F1 - Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

C. Experimental Results

Implementation is carried out in Python language in Google Colab Platform. Artificial Neural Network is created with four hidden layers. The activation function used is ‘ReLU’. Table 1 and Figure 7 shows the F1-score, precision, recall and accuracy of ANN with varying Epochs and Batch size. Epochs refers to the number of iterations and Batch size refers to the varying number of input instances.

Epochs	Batch Size	Precision	Recall	F1-Score	Accuracy
250	15	0.68	0.68	0.69	73.9
250	30	0.74	0.75	0.74	80.78
250	45	0.74	0.75	0.74	82.62
250	60	0.76	0.77	0.76	83.06
500	45	0.78	0.79	0.78	84.64

Table 1. Performance of ANN with varying Epochs & Batch size

Comparison of Precision, Recall & F1-Score with varying Batch Size

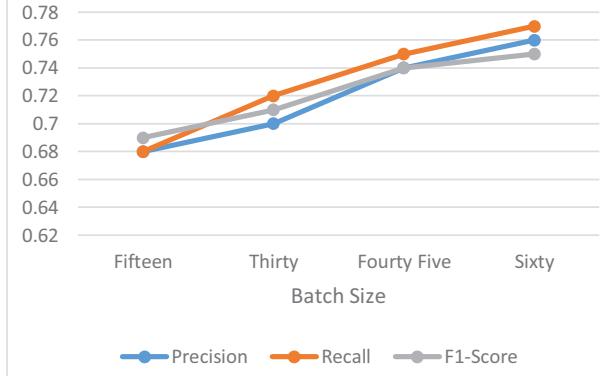


Fig. 7. Performance of ANN with Epoch value 250 & varying Batch size

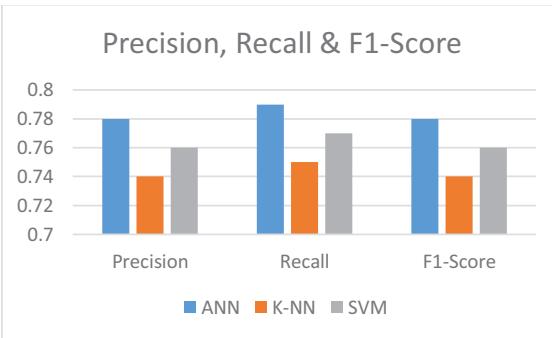


Fig. 8.Precision, Recall & F1-Score

Table 2 and Figure 8 shows the comparison of the precision, recall, F1-Score and Accuracy of the ANN, K-NN and SVM Classifiers.

Classifier	Precision	Recall	F1-Score	Accuracy
ANN	0.78	0.79	0.78	84.64
K-NN	0.74	0.75	0.74	76.34
SVM	0.76	0.77	0.76	81.65

Table 2. Performance of ANN, K-NN and SVM.

V. CONCLUSION

Diabetes Mellitus is increasing rapidly among the common population which may result in other serious health issues. It is important to diagnose the disease at earliest and to give proper treatment. In this Paper, the input instances are extracted from the UCI Machine Learning repository and popular supervised learning algorithms that use labelled data as input namely SVM, ANN and K-NN are used in the prediction of Diabetes. Different techniques give different accuracy on the same dataset. The evaluation metrics such as accuracy, precision, recall and F1-score of all three algorithms are analyzed and compared. From the implementation result values, it is found that ANN outperforms the other two classifiers taken for evaluation. Further the efficiency of the classification of ANN relies on factors like number of layers used in hidden layers, number of epochs, number of neurons, batch size etc., If these parameters are carefully chosen, ANN can give better Classification results. The work can be extended by finding the possibility of occurrence of Diabetes in a Non-Diabetic Person.

REFERENCES

- [1] Y. K. Qawqzeh, A. S. Bajahzar, M. Jemmali, M. M. Otoom, and A. -aljaoui, "Classification of diabetes using photoplethysmogram (PPG) waveform analysis: logistic regression modeling," *BioMed Research International*, vol. 2020, Article ID 3764653, 6 pages, 2020.
- [2] G. A. Pethunachiyar, "Classification of diabetes patients using kernel based support vector machines," in *Proceeding of the 2020 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–4, IEEE, Coimbatore, India, January 2020.
- [3] S. Gupta, H. K. Verma, and D. Bhardwaj, "Classification of diabetes using Naive Bayes and support vector machine as a technique," *Operations Management and Systems Engineering*, Springer, Singapore, pp. 365–376, 2021.
- [4] D. K. Choubey, M. Kumar, V. Shukla, S. Tripathi, and V. K. Dhandhania, "Comparative analysis of classification methods with PCA and LDA for diabetes," *Current Diabetes Reviews*, vol. 16, no. 8, pp. 833–850, 2020.
- [5] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, "Classification and prediction of diabetes disease using machine learning paradigm," *Health Information Science and Systems*, vol. 8, no. 1, pp. 7–14, 2020.
- [6] R. Ahuja, S. C. Sharma, and M. Ali, "A diabetic disease prediction model based on classification algorithms," *Annals of Emerging Technologies in Computing*, vol. 3, no. 3, pp. 44–52, 2019.
- [7] S. K. Mohapatra, J. K. Swain, and M. N. Mohanty, "Detection of diabetes using multilayer perceptron," in *Proceeding of the International Conference on Intelligent Computing and Applications*, pp. 109–116, Springer, Ghaziabad, India, December 2019.
- [8] N. Singh and P. Singh, "Stacking-based multi-objective evolutionary ensemble framework for prediction of diabetes mellitus," *Biocybernetics and Biomedical Engineering*, vol. 40,no. 1, pp. 1–22, 2020.
- [9] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," *International Journal of Cognitive Computing in Engineering*, vol. 2, 2021.
- [10] M. M. F. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra, "Likelihood prediction of diabetes at early stage using data mining techniques," in *Computer Vision and Machine Intelligence in Medical Image Analysis*, pp. 113–125, Springer,Singapore, 2020.
- [11] S. Malik, S. Harous, and H. E. Sayed, "Comparative analysis of machine learning algorithms for early prediction of diabetes mellitus in women," in *Proceedings of the International Symposium on Modelling and Implementation of Complex Systems*, pp. 95–106, Springer, Batna, Algeria, October 2020.
- [12] A. Hussain and S. Naaz, "Prediction of diabetes mellitus: comparative study of various machine learning models," in *Proceeding of the International Conference on Innovative Computing and Communications*, pp. 103–115, Springer, Delhi, India, January 2021.
- [13] Choudhury A, Gupta D (2019) "A survey on medical diagnosis of diabetes using machine learning techniques". In: Kalita J, Balas VE, Borah S, Pradhan R (eds) Recent developments in machine learning and data analytics. Advances in intelligent systems and computing, vol 740. Springer, Singapore, pp 67–78. https://doi.org/10.1007/978-981-13-1280-9_6
- [14] Dalakleidi KV, Zarkogianni K, Karamanos VG, Thanopoulou AC, Nikita KS (2013) "A hybrid genetic algorithm for the selection of the critical features for risk prediction of cardiovascular complications in Type 2 Diabetes patients." In: Abstracts of the 13th IEEE international conference on BioInformatics and BioEngineering, Chania, 10-13 November 2013. <https://doi.org/10.1109/BIBE.2013.6701620>
- [15] Harris MI, Klein R, Welborn TA, Knutiman MW (1992) "Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. *Diabetes Care*" 15(7):815–819. <https://doi.org/10.2337/diacare.15.7.815>
- [16] Ameena RR, Ashadevi B (2020) "Predictive analysis of diabetic women patients using R. In: Peter JD, Fernandes SL (eds) Systems simulation and modeling for cloud computing and big data applications." Elsevier Inc., Amsterdam. <https://doi.org/10.1016/B978-0-12-819779-0-00006-X>
- [17] Daanouni O, Cherradi B, Tmiri A (2019) "Predicting diabetes diseases using mixed data and supervised machine learning algorithms. In: Abstracts of the 4th international conference on smart city applications", ACM, Casablanca,2019. <https://doi.org/10.1145/3368756.3369072>
- [18] Sisodia D, Sisodia DS (2018) "Prediction of diabetes using classification algorithms." *Procedia Comput Sci* 132:1578–1585. <https://doi.org/10.1016/j.procs.2018.05.122>
- [19] Ahuja R, Sharma SC, Ali M (2019) A diabetic disease prediction model based on classification algorithms. *Ann Emerg Technol Comput* 3(3):44–52. <https://doi.org/10.33166/AETiC.2019.03.005>
- [20] Alehegn M, Joshi RR, Mulay P (2019) Diabetes analysis and prediction using random forest, KNN, Naïve Bayes, and J48: an ensemble approach. *Int J SciTechnol Res* 8(9):1346–1354
- [21] S. N. Bushra and G. Shobana "A Survey on Deep Convolutional Generative Adversarial Neural Network (DCGAN) for Detection of Covid-19 using Chest X-ray/CT-Scan" 3rd Scopus-Indexed IEEE International Conference on Intelligent Sustainable Systems (ICISS 2020). Dec 3, 2020 - Dec 5, 2020.
- [22] Mohan, Narendra & Jain, Vinod. (2020). "Performance Analysis of Support Vector Machine in Diabetes Prediction". 1-3. 10.1109/ICCEA49313.2020.9297411.
- [23] Pethunachiyar, G.. (2020). "Classification Of Diabetes Patients Using Kernel Based Support Vector Machines." 1-4. 10.1109/ICCCI48352.2020.9104185.
- [24] Al-Shayea, Qeethara. (2011). "Artificial Neural Networks in Medical Diagnosis". *Int J Comput Sci Issues*. 8. 150-154.
- [25] Mailagaha Kumbara, Mahinda & Luukka, Pasi & Collan, Mikael. (2020). "A new fuzzy k-nearest neighbor classifier based on the Bonferroni mean. Pattern Recognition Letters". 140. 172-178. 10.1016/j.patrec.2020.10.005.

Categorization of ECG Signal Using CNN Algorithm

Md Tarique Anis
Department of Electrical Engineering
National Institute of Technology Hamirpur
Hamirpur, India
20mee108@nith.ac.in

Dr Veena Sharma
Department of Electrical Engineering
National Institute of Technology Hamirpur
Hamirpur, India
veena@nith.ac.in

Abstract— It is widely known that heart disease is among the most significant cause of loss of life worldwide. Therefore, early detection of heart diseases is important to reduce the rising death rate. An electrocardiogram can detect many types of heart diseases including abnormal heart rhythms. We propose here a method for categorizing heart diseases entirely focused on ECG by using a machine learning method, known as Convolution Neural Networks (CNNs) into five categories as per the Association for Advancement of Medical Instrumentation (AAMI) standards. The outcomes conveys that the suggested methodology outperforms several previously available methods for ECG signal classification in respect of the classification accuracy and computing efficiency.

Keywords— Convolution Neural Networks (CNNs), ECG, Electrocardiogram, AAMI Standard.

I. INTRODUCTION

According to many health organizations including the WHO the cardiovascular illnesses is one of the leading causes of mortality worldwide. An electrocardiogram (ECG) is one of the most commonly used method used for getting the electrical activity generated by the heart. It is commonly used because of its non-invasive method. It is commonly employed in the detection of cardiovascular diseases. The term Arrhythmia refers to irregularities in the cardiovascular functions such as the rhythm, rate or the conduction of electrical signals through the heart and is the most commonly referred as cardiac diseases [1]. And these cardiac abnormalities can affect cardiac electrical activity, which can be recognized by analyzing an ECG waveform, composed of various electrical signals that are linked with heart activity and can provide important information about a patient's heart condition.

There are mainly three types of ECG Signals:

- Resting ECG: This ECG is carried out when the subject is lying down comfortably.
- Stress or Exercise ECG: This ECG is carried out when the subject is using a treadmill or doing exercise on an exercise bike.
- Ambulatory ECG or Holter ECG: In this type of ECG a small portable device consisting of electrodes is

wrapped around the waist which can monitor the heartbeat for one or more days

To get the most accurate diagnosis, an ECG using 10 electrodes capturing 12 leads (signals) is performed. Each lead examines the electrical activity of the cardiovascular system in a different way. 12 leads must be used to record an accurate result. Each ECG consist of five waves, which correspond to different phases of the heart's activity: P, Q, R, S, and T. In ECG waveform the first positive deflection is known as P wave. Depolarization of the ventricles is represented by the QRS complex, but the fact is that QRS complex may not always display all three waves, and its normal duration is between 0.08 seconds to 0.10 seconds i.e 80 to 100 milliseconds.

Often, for doctors it is difficult to examine the lengthy ECG recordings in a relatively short time frame and observing the alterations caused by the abnormalities of the heart and the cardiovascular system, hence there is a need for automatic classification of ECG signals.

Automation of ECG signal categorization is a difficult task to solve for many points. Primary ECG patterns from different patients can have morphological and temporal differences in their waveforms. Distinct patients with different cardiac beats may have similar ECG waveforms, while the same patient may have different ECG waveforms recorded at different time interval. The other factor involved in ECG signal classification is heart beat variability. Heart rate is affected by conditions such as stress, excitement, and exercise, and can result in difference in the ECG signal features, also there is no optimal classification rule for ECG signals.[1] ECG arrhythmia classification also faces the challenge of developing the classifier capable of classifying arrhythmia in real time.

Automatic heart rate classification has been previously reported by many researchers using different functions and different classification methods to represent ECG signal [2]. Many classification and detection algorithms have been developed based on different techniques like artificial neural networks [3], hidden Markov models [4] and many more. But the main problem associated with the above classification model is that

they perform good on training data but when exposed to ECG signal of different patient performance deteriorates. The extraction of features is an important step in ECG categorization. Various preprocessing and feature extraction strategies for ECG classification have been published by a number of researchers. Due to the diversity and individuality of ECG data for same and different patient, such a self-customized, and fixed feature might not work well for all individuals. To overcome the above stated limitations and shortcomings, we present a classification system based on deep learning using the famous convolutional neural networks (CNNs).

The work in this paper is arranged as follows: Previous work is reviewed in section II. Section III discusses the openly accessible dataset employed in this study and gives a comprehensive summary of ECG data, it also includes how the raw data is represented and how the test and training datasets are created for the suggested classification system. The performance of the suggested methodology for the classification is evaluated in Section IV, and the outcomes are compared to other available methods. And at the last in Section V, the paper's conclusions are presented.

II. BACKGROUND AND RELATED WORK

The ECG waveform represents the electrical activity of the cardiovascular system and can aid in the detection of many heart rate irregularities and rhythm. As a result, they serve a crucial role in the detection of cardiac diseases. Preprocessing, feature extraction, and classification are the three essential and basic steps in ECG pattern categorization or classification. There are number of studies related to ECG classification where the size of the data is not very large that did not use any big data tool. But on the other side there are many studies that uses large dataset and big data techniques.

In Indonesia, cardiovascular diseases cause a high rate of mortality. To minimize it, a tele-ECG system using Hadoop framework was developed for detecting and monitoring heart diseases early. Decision trees (DT) and random forests (RF) are used to classify the ECG data. This is the first model to use big data tools for heartbeats classification. The base for the system was 4 nodes cluster computer. The server could serve 60 requests simultaneously with an accuracy of 97.14% and 98.92% respectively [5].

In [6], researchers have used deep neural networks (DNNs) to classify heartbeats. The accuracy of classifying has been demonstrated at 99%, but only two types of heartbeats were classified (Abnormal and Normal), and the size of the dataset was approximately 85,000. And in [7], a patient-adaptable algorithm to classify the ECG heartbeats is presented. An algorithm for classifying RR intervals and R peaks was carried out by A.Vishwa et al. [8], using the MIT-BIH arrhythmia dataset and the Normal sinus rhythm dataset. He applied Feed Forward Neural Network (FFNN) algorithm, along with error back technique for classification. D.Patra et al in [9], used a 3-layer Feed Forward Neural Network (FFNN) along with back

propagation technique as a classification algorithm, database used in the above stated technique is MIT-BIH arrhythmia database, and for feature normalization and reduction, the zero mean and FCM are used respectively.

In [10], ECG signals were classified using probabilistic neural networks into normal and arrhythmia categories on the MIT-BIH arrhythmia database and the classification accuracy was found to be 96.5%.

In [11], a classification of ECG images was conducted using artificial neural networks. In this paper artificial multilayer perceptron-type neural networks and haar-like feature extraction algorithms are used to maximize the accuracy.

In [21], to classify different cardiac arrhythmia four steps were used, it starts with the pre-processing of the data, then conversion of the 1-D ECG data into 2-D scalogram image using the Continuous Wavelet Transform (CWT) technique, and finally the classification of the arrhythmia into different types using the Transferred Deep Convolutional Network

From the literature conferred, it is clear that artificial neural networks have emerged as the most popular and extensively used algorithm for the classification.

The recent advancement in the deep learning has made it emerge as powerful learning capacity tool and also has shown excellent outcomes in many fields like object recognition [12], image processing and classification [13], biomedical image processing [14], and also in time series data [15]. In another study [16], 1-D convolutional neural networks (CNNs) were utilized to classify patient specific real-time ECG signals. MIT-BIH arrhythmia database was evaluated and the classifier has shown notable classification performance when compared with other available methods for classification of anomaly beats.

However, due to the difficulties outlined in section 1, none of the proposed techniques have functioned successfully, most of these methods exhibit inconsistent results when attempting to categorize an ECG signal from a novel patient. The trail of most of these methods was conducted on a small ECG database, these methods also do not satisfy the AAMI criteria for class labelling.

To address the aforementioned limitations and shortcomings, we propose an ECG heart beat categorization approach based upon the convolutional neural networks in this study.

III. DATA PREPROCESSING

A. Dataset

All experiments presented in this paper were conducted with an ECG dataset obtained from the MIT-BIH arrhythmia database. This dataset comprises of 48 recordings of two channel ambulatory ECGs, all consisting of 30 minutes of selected segments from 24 hours of recordings of 48 different subjects. 0.1-100Hz band pass filters are used to pre-process the ECG signals, and 360Hz sampling is performed. The AAMI suggestion for class labelling is used in the database. According to AAMI every single beat should be categorized into one of the five types: N (normal beats),

TABLE 1: Class description of ECG Signal using AAMI standards [20]

AAMI Classes →	Normal beat (N)	Supraventricular ectopic beat (S)	Ventricular ectopic beat (V)	Fusion beat (F)	Unknown beat (Q)			
MIT-BIH Arrhythmia Classes	Normal beat (N)	Atrial premature beat (A)	Premature ventricular contraction (V)	Fusion of ventricular and normal beat (F)	Paced beat (/)			
	Left bundle branch block beat (L)	Aberrated atrial premature beat (a)	Ventricular escape beat (E)		Fusion of paced and normal beat (f)			
	Right bundle branch block beat (R)	Nodal (junctional) premature beat (J)			Unclassified beat (Q)			
	Atrial escape beat (e)	Supraventricular premature beat (S)						
	Nodal (junctional) escape beat (j)							

S (supraventricular ectopic beats), V (ventricular ectopic beats), and F (fusion beats), and Q (unclassified beats). We used signals from modified-lead II for all records, and different labels were used for localization of ECG beats. Table 1 includes an ECG class description based on AAMI guidelines [20]. A total of 120 data samples are extracted from both right and left sides of the R-peak. Before feeding the data into the input of CNN layer, the data is down-sampled such that each beat of raw data is represented by 128 samples.

B. Proposed Approach

An effective ECG beat categorization system based on convolutional neural networks (CNNs) is presented in this research. The Convolutional Neural Networks popularly called as a feature learner [1], and have a lot of potential for extracting useful features from input data. Fig.1 provides an overview of proposed methodology. As stated in section II essential and basic steps in ECG beat classification are (a) preprocessing, (b) feature extraction, and (c) classification. In order to obtain a beat representation for CNN, equal numbers of samples from left and right sides of the R-peak are taken and are down-sampled for further processing.

There are two parts in the convolutional neural networks. The initial half is a feature extractor that has the capability to learn features automatically from raw input, while the other part is a fully linked multi-layer perceptron (MLP) that executes classification with the features acquired from the initial half. Further, convolutional and pooling layers are present inside the feature extractor as the main building block [1]. Building block of CNN are the convolutional layer comprised of a series of

filters (or kernels), whose parameters are trained over the training process. These filters have a smaller height and weight than the input data.

Every filter is convolved well with given data to produce a neuron-based activation map. By utilizing the interconnection of the convolution layers, the network filter learns and respond maximally to a specific area of the given input data, and thereby take the advantage of the correlation among the input functions. In most of the cases, the pooling layer is sandwiched in the middle of the two convolutional layers. By downsampling the representation, the pooling layer bring down the number of variables and also the calculation burden.

There are two types of pooling: maximum and average. Maximum or max pooling is frequently used because it performs well. In this paper max pooling is used which gets the highest value in a group of nearby inputs.

A convolution is performed by convolving the input (I) with the number of filters as follows:

$$y_k = I * P_k + b_k, \quad k = 1, 2, 3, \dots, Z \quad - (1)$$

Where Z is the number of filters, y_k is the output corresponding to k^{th} convolution filter, P_k is the weight of the k^{th} filter. And b_k is the k^{th} bias.

To extract high-level features, more than one convolutional and pooling layers can be layered on each other. In addition to convolution, batch normalization is used after convolution to decrease the effect of initialization and speeds up the net training. To boost the network's nonlinear properties, nonlinear activation functions are applied element by element. The most widely used nonlinear activation function is ReLU. In our paper

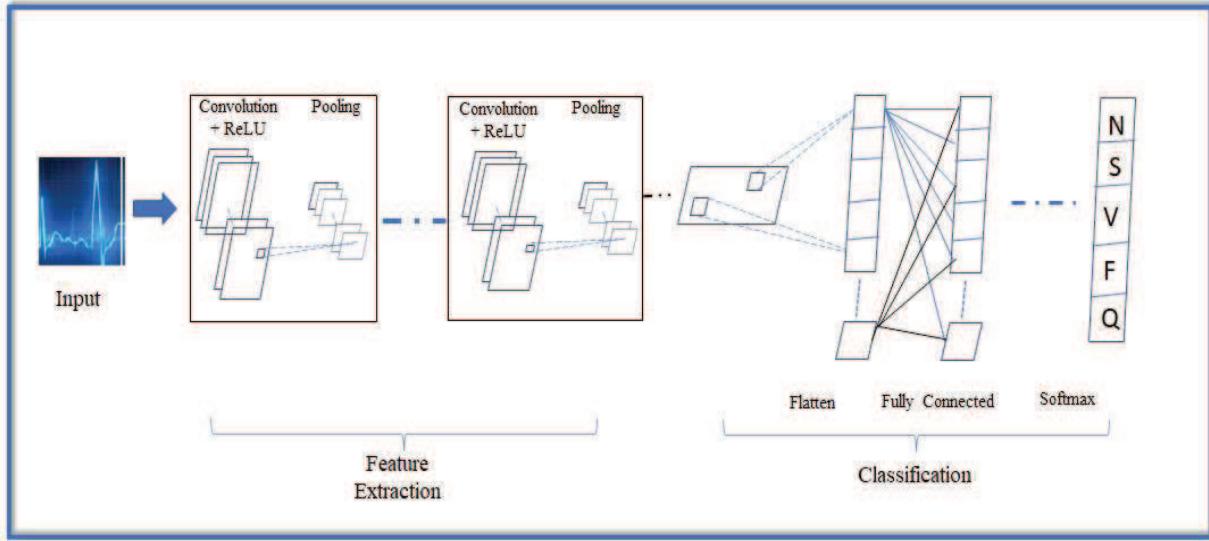


Fig.1: Overview of the presented methodology

ReLU is used as the non-linear activation function. After the last convolutional layer, the softmax layer is being used to produce the probability distribution of the classification performance for each pixel.

IV. RESULTS

A. Experimental Setup

As discussed earlier in section III, we represented every beat as 128 samples focused around the R-Peak for the analysis of proposed ECG beat categorization system. And in our research, we employed an uncomplicated 1-D CNN consisting of four convolutional layers, four pooling layers, one softmax layer and also a flatten layer. For the first two convolutional layers the kernel size was 3 and subsampling factor for pooling layer was set to 3. ReLU is used as non-linear activation function. Dropout layer is used to prevent the model from overfitting with the parameter value of 0.5.

B. Performance Evaluation

To categorize ECG beats, data from the MIT-BIH arrhythmia database is used which consists of ECG data of 48 individuals. Based on the recommendations of the AAMI, these 48 records with 100389 beats have been classified into five types of heartbeat (i.e N-type, S-type, V-type, F-type, Q-type). From the dataset 70% of the data is selected at random from each type for the training purpose and remaining beats were used to evaluate test patterns.

As a validation dataset, 8,000 beats were picked at random from the test dataset. The proposed approach in the paper has an overall accuracy of 98.4% for the validation set and

96.6% for the test set. Fig.2 shows the confusion matrix of ECG beat categorization system.

From the confusion matrix the data can be seen as follows: The overall testing accuracy is 96.6%. And the accuracy for classifying the N-type beat is 96.71, for S type beat is 93.31, for V-type beat is 96.53, and for F and Q type is 84.65 and 99.47 respectively.

The proposed algorithm in this paper was implemented without the use of graphics processing units (GPU) in Python on a computer with a i5 processor running at 2.42 GHz and 8 GB of memory. And the proposed model has a significantly low computational cost. In table II, we compare our algorithm to three other existing ones.

TABLE III: Comparison of model with other techniques

	Methods			
	M. Zubair et al (2016) [1]	N.P. Joshi et al [19]	V.Srivastava et al (2013) [18]	Proposed↓
Accuracy-	91.7%	85%	85%	96.6

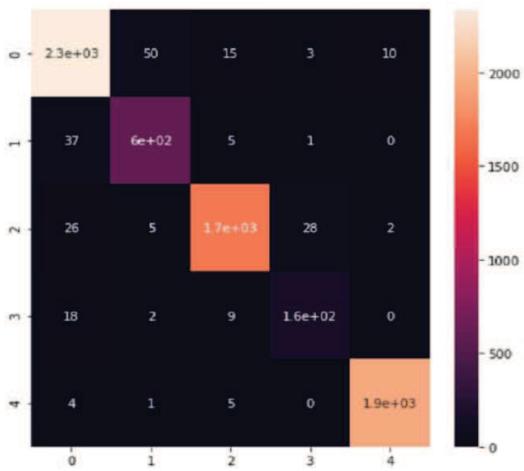


Fig.2: Confusion Matrix

V. CONCLUSION

This paper proposes a classifier for ECG signal based on convolutional neural networks (CNNs), that is able to extract and learn features of the provided raw ECG data as the input. ECG data from MIT-BIH database is used as input to evaluate the performance of the proposed system. And as per AAMI standards each beat was classified as one of the five beat types. As demonstrated in the study, the suggested model has impressive accuracy in classification and high computing proficiency.

Future work will focus on improving performance by examining other features of the ECG signal by using other improved deep learning-based ECG classification algorithms.

REFERENCES

- [1] Muhammad Zubair, Jinsul Kim and Changwoo Yoon, "An automated ECG beat classification using Convolutional Neural Networks," International Conference on IT Convergence and Security (ICITCS), 2016, pp. 1-5
- [2] Christov I, Jekova I, Bortolan G. "Premature ventricular contraction classification by the kth nearest-neighbours rule." *Physiol Meas*. 2005;26(1):123
- [3] Y. H. Hu et al., "Applications of artificial neural networks for ECG signal detection and classification," *J. Electrocardiol.*, vol. 26, pp. 66–73, 1994.
- [4] D. A Coast et al., " An approach to cardia arrhythmia analysis using hidden Morkow models," *IEEE Trans. Biomed. Eng.*, vol. 37, no.9, pp. 826-836, Sep. 1990
- [5] Ma'sum M.A, Jatmiko W, Suhartanto H. Enhanced tele ECG system using hadoop framework to deal with big data processing. In: 2016 international workshop on Big Data and information security (IWBS). New York: IEEE; 2016. p. 121–6.
- [6] Celesti F, Celesti A, Carnevale L, Galletta A, Campo S, Romano A, Bramanti P, Villari M. Big data analytics in genomics: the point on deep learning solutions. In: 2017 IEEE symposium on computers and communications (ISCC). New York: IEEE; 2017. p. 306–9.
- [7] Llamedo et al., "A automatic patient-adapted ECG heartbeat classifier allowing expert assistance", *IEEE Transaction on Biomed. Eng.*, vol. 59, no. 8, pp. 2312-2320, 2012.
- [8] A. Vishwa, M. K. Lal, S. Dixit, and P. Vardwaj, "Clasification of arrhythmic ECG data using machine learning techniques," *Int. J. of Interactive Multimedia and Artificial Intell.*, vol. 1, no. 4, pp. 68-71, 2011.
- [9] D. Patra, M. K. Das, and S. Pradhan "Integration of FCM, PCA and neural networks for classification of ECG arrhythmias," *IAENG Int. J. of Comput. Sci.*, vol. 36, no. 3, pp. 24-62, 2010.
- [10] M. Vijayanan, V. Rathikarani, and P. Dhanalakshmi, "Automatic Classification of ECG Signal for Heart Disease Diagnosis using morphological features," *Int. J. of Comput. Sci. and Eng. Technology (IJCSET)*, vol. 5, no. 4, pp. 449-455, 2014.
- [11] B. Mohamed et al., "ECG image classification in real time based on the harr-like feature and artificial neural networks", In procedia Computer Science vol. 73, pp. 32-39, 2015.
- [12] J. Bai , Y. Wu , J. Zhang , F. Chen , "Subset based deep learning for RGBD object recognition," *Neurocomputing* vol.165, pp.280–292, 2015.
- [13] W M. Hayat , M. Bennamoun , S. An , "Deep reconstruction models for image set classification," *IEEE Trans. Pattern Anal. Mach. Intell.* Vol.37, pp. 713-727, 2015.
- [14] T. Brosch , R. Tam , Efficient training of convolutional deep belief networks in the frequency domain for application to high-resolution 2D and 3D images, *Neural Comput.* Vol.27, pp. 211–227, 2015
- [15] J. Deng , Z. Zhang , F. Eyben , B. Schuller, "Autoencoder-based unsupervised domain adaptation for speech emotion recognition, *IEEE Signal Process. Lett.*" Vol. 21,pp. 1068–1072, 2014.
- [16] S. Kiranyaz et al., " Real time patient-specific ECG classification by 1-D convolutional neural networks", *IEEE Trans. On Biomed. Eng.*, vol. 63, no. 3, 2016.
- [17] V. Srivastava and D. Prasad, "Dwt-Based Feature Extraction from, ecg Signal," *American J. of Eng. Research (AJER)*, vol. 2, no. 3, pp. 44- 50, 2013.
- [18] N. P. Joshi and P. S. Topannavar, "Support vector machine based heartbeat classification," *Proc. of 4th IRF Int. Conf.*, pp. 140-144, 2014.
- [19] X. Tang and L. Shu, "Classification of Electrocardiogram Signals with RS and Quantum Neural Networks," *Int. J. of Multimedia and Ubiquitous Eng.*, vol. 9, no. 2, pp. 363-372,2014.
- [20] Pandey, S.K., Janghel, R.R., Dev, A.V. et al. "Automated arrhythmia detection from electrocardiogram signal using stacked restricted Boltzmann machine model." *SN Appl. Sci.* 3, 624 (2021)
- [21] P. Giriprasad Gaddam et al "Automatic Classification of Cardiac Arrhythmias based on ECG Signals Using Transferred Deep Learning Convolution Neural Network" 2021 *J. Phys.: Conf. Ser.* 2089 012058

Path Planning of Intelligent UAVs using Computational Geometric Techniques

Akshya. J
School of Computing,
SASTRA Deemed to be University,
Thanjavur, India
akshya@sastra.ac.in

Prudhvi Gadupudi
School of Computing,
SASTRA Deemed to be University,
Thanjavur, India
gadupudiprudhvi@gmail.com

P.L.K.Priyadarshini
School of Computing
SASTRA Deemed to be University
Thanjavur, India
priya.ayyagari@it.sastra.edu

Karthik Paladugula
School of Computing
SASTRA Deemed to be University
Thanjavur, India
karthik51199@gmail.com

Abstract— In recent years, Unmanned Aerial Vehicles, commonly referred to as UAVs, are heavily used in various real-time scenarios without involving human interaction. Also, they are among the most extensively researched topics. The profit of using UAVs has also come up with different challenges. One of the most important among them is the Coverage Path Planning problem, generally abbreviated as CPP. Computation of path also involves the consideration of various static and dynamic obstacles. While considering geographic areas as polygons, path planning of UAVs becomes difficult for concave polygons. To obtain an optimal path in a concave polygon, we have proposed the Field of View-based Sweep-Line method(FoVSL) technique. This proposed work has enhanced the existing Sweep-Line method for finding the optimal path in a polygonal area. As the solution can be obtained with ease when the polygons are convex, the use of the convex polygonal decomposition technique is adapted with the proposed method. The inclusion of the Boustrophedon method reduced the total number of turns taken by the UAV. Comparative analysis with the existing technique proves that the proposed method minimizes the total path to be traveled with a reduction in the number of turns taken by the UAV resulting in overall minimization of the time consumed for the entire traversal.

Keywords—Unmanned Aerial Vehicles, Travelling Salesperson Problem, Coverage Path Planning, Boustrophedon Method

I. INTRODUCTION

Autonomous UAVs are deployed in various applications where the reach of humans remains impossible or would require more manpower [1]. The use of UAVs is done in many surveillance applications to ease the work. Some of the applications include military, healthcare, smart city, and a lot more [2]. With the deployment of UAVs, there is an ease of work and has also reduced the use of manpower to a greater extent. With the deployment of UAVs, there exist several challenges some of them being collision, battery life, etc [3]. Coverage Path Planning(CPP) [4] is a problem of significance for a UAV in any real-time surveillance application where an optimal path needs to be found out such that it covers all the given set of locations [5]. To obtain this optimal path, we require computationally efficient algorithms. In most of the surveillance applications, UAV's starting and ending points are the same. It should visit all the given points as shown in Figure 1. In

this, it could be seen that the UAV starts from a starting point, covers all the other points, and reaches back to the same point from where it started. Considering any path planning algorithm, the vital aim is to minimize the traversal time of UAVs without any collision. The path taken by the UAV needs to be optimal and should also consider other constraints such as static and dynamic obstacles, maneuverability time, wind direction, and number of turns.

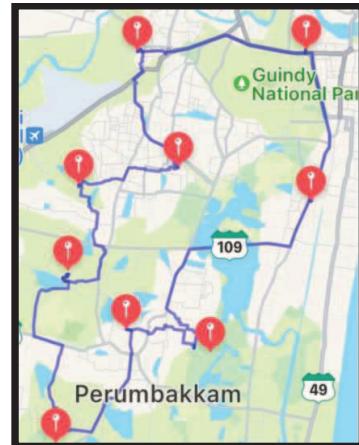


Figure 1. Real-time Surveillance application

In this paper, we have proposed the FoVSL technique to enhance the existing Sweep-Line technique. Comparative analysis is done to prove the efficiency of the proposed method. The rest of the paper examines various existing works done on computing optimal paths for UAV, proposed methodology used for obtaining an optimal path, results obtained while evaluating the proposed method with the existing methods, and finally concludes by proposing future works.

II. RELATED WORKS

As the deployment of UAVs is increasing day by day, numerous researchers are focusing on providing various solutions to the problems arising while deploying UAVs in a real-time environment. This section summarizes the existing works on path planning of UAVs and also

highlights some of the research gaps identified. Though the deployment of UAVs has been increased in recent times, there exist various constraints that need to be focused on before using the UAV in a real-time surveillance application. Some of the constraints include the computation of optimal path, traversal time, battery life of UAV, and collision avoidance. Some intelligent UAVs are designed to satisfy one objective while others try to solve multiple objectives in a given time. Coverage Path Planning(CPP) is one of the critical problems that need to be addressed in particular for making the deployment of UAV more effective and autonomous for any given real-time application.

To solve CPP problems effectively numerous researchers have come up with various contributions. Focusing on partitioning the given geographic area, J. Hu et.al.[6] has adopted Voronoi-based partitioning algorithm with non-overlapping regions. The major research gap identified in this work includes not considering the height of the UAV while partitioning the area. In [7], the author has chosen the method of Delaunay triangulation as area partitioning technique whereas in [8], decomposition was done such that the area was decomposed into equal-sized partitions. In the former, the author considered only a constant FoV of 100*900 meters whereas, in the latter, only equal-sized partitions were taken into consideration.

Looking at the research works carried on identifying the various coverage points for finding an optimal path within the area partitioned, a coverage path was built on regular shaped areas by considering the vertex of the area as the point of coverage [9]. The method failed to consider both static and dynamic obstacles. M.Coombes et.al. [10] decomposed the area into convex polygons while considering obstacles and wind direction, but failed to consider the FoV of the UAV. Adopting to Grid-based partitioning technique, the author in [11] has covered the entire area by choosing grid points but the grids had a constant size of 1*1 square units.

Research works on optimal path planning of UAVs using the identified coverage points include the use of Graph Neural Network where the entire problem was posed as TSP. The major research gap identified in this paper included the use of only symmetric graphs [12]. Considering [13], the use of the Artificial Bee Colony algorithm was done to find the optimal path. Choice of this particular algorithm was made over the others as it avoided being stuck in local minima. However, the algorithm failed to cover the entire geographic area. H. Azpúrrua et. al. [14] built an effective coverage trajectory for the robots but failed to consider the path in presence of obstacles which is very essential while deploying UAVs in a real-time environment.

III. METHODOLOGY

The problem under consideration in this research work is to find an optimal path for a UAV for any given geographic area in the presence of static obstacles. The methods used for obtaining this path are discussed in this section.

A. Boustrophedon Method

In this method, the computation of the path is done in such a way that it is perpendicular to the width of the

polygon [15]. This in turn reduces the number of turns to be taken by the UAV as shown in Fig. 2. The use of the Boustrophedon method can be applied only on convex polygons. But in real-time surveillance applications, UAVs need to traverse through concave polygons and hence computation of optimal paths using the Boustrophedon method becomes difficult. To overcome this challenge, we have adapted the technique of Convex Polygonal Decomposition.

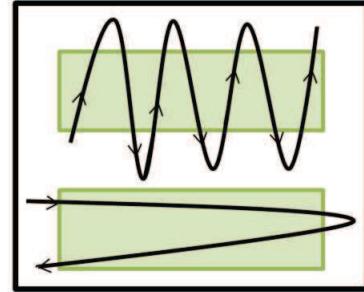


Figure 2. Use of Boustrophedon method

B. Convex Polygonal Decomposition

In this method, the concave polygons are decomposed in the presence of static obstacles so that implementation of the Boustrophedon method is possible. The key idea behind this method is to decompose the concave polygons to form convex sub-polygons [13]. To obtain this, interior extension of edges of the static obstacles are done till they hit the boundaries of the polygon. Once the edges are extended, the merging of obtained sub-polygons is done as long as the merged polygons are convex as shown in Fig. 3.

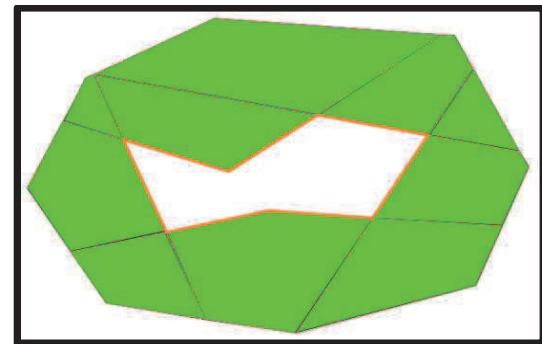


Figure 3. Convex Polygonal Decomposition

C. Sweep-Line Method

In this method, the UAV traverses through the geographic area in a to and fro fashion such that it covers the entire area [16]. The traversal is based on the FoV of the UAV. In this method, the FoV is considered only while traversing in a straight line but is not considered while making turns. The depiction of the Sweep-Line method over a given geographic area is shown in Fig. 4.

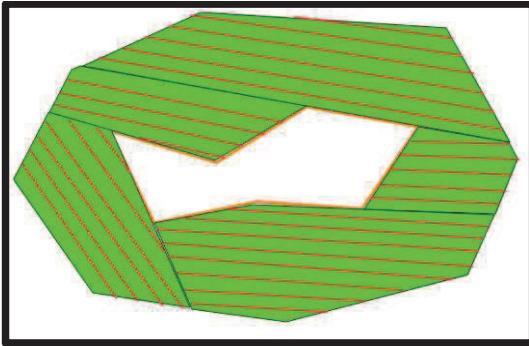


Figure 4. Sweep-Line Method

D. Proposed Algorithm - FoV based Sweep Line Method

In this research work, we have proposed an algorithm namely, FoVSL. This algorithm finds an optimal path throughout the given geographic area by making use of the above-stated methods in this section. The novelty of the proposed algorithm is that, while implementing the traditional Sweep-Line method, the Field of View (FoV) [17] is not considered when the UAV makes a turn. FoV is considered only along the paths covered while traversing in straight lines. In the proposed method, the FoV of the UAV is considered even while making turns. Apart from this, coverage points are also identified for each sub-polygon. Each sub-polygon needs to have an entry and an exit point and these points are identified by adapting the Boustrophedon method. With the use of this method, there is a minimization in the total path covered by the UAV and hence, reduces the overall traversal time. Fig. 5, depicts the implementation of the proposed algorithm over the concave polygons in the presence of static obstacles.

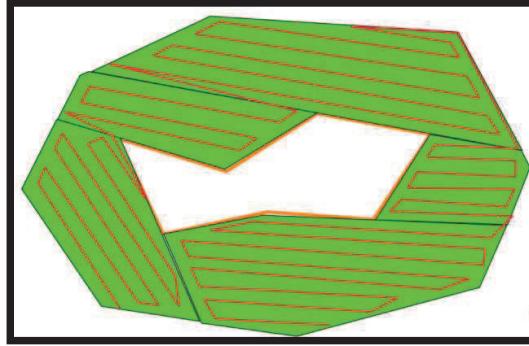


Figure 5. Computation of near-optimal path

IV. EXPERIMENTAL RESULTS

The experimental results for obtaining the optimal path for the given geographic area were done using QGIS 3.10.0 with GRASS 7.8.0. Initially, the flying height of the UAV is set such that the traversal is based on the Field of View(FoV) of the UAV. The proposed method is implemented on a map by depicting a UAV such that it gives a better understanding of the proposed method. It is noted that the starting and the ending points of the optimal

path are the same. The UAV reaches the same position after the entire traversal from where it started. The method can be applied to both, polygons without any static obstacle as shown in Fig. 6 and polygons with static obstacles as shown in Fig. 7.

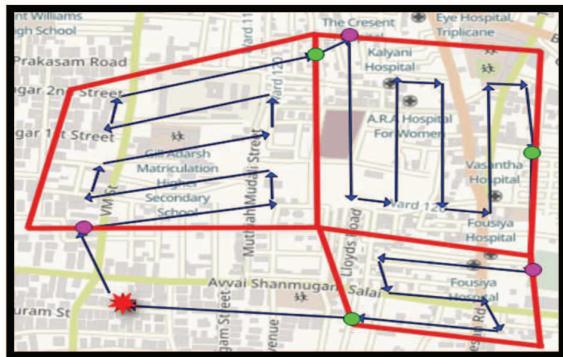


Figure 6. FoVSL method without obstacle



Figure 7. FoVSL method with obstacle

To verify the effectiveness of our algorithm, it's been compared with the existing algorithm Sweep-Line technique ad is tabulated in Table I and Table II respectively.

TABLE I. EXPERIMENTAL RESULTS FOR EXISTING - SWEEP-LINE

Polygon	Existing technique (Sweep-Line)		
	Tour Length (in km)	Total number of turns	Execution Time (in s)
P1	1.4	10	6
P2	1.7	15	14
P3	1.5	12	18
P4	1.6	8	9
P5	2.5	21	13
P6	3.9	23	19
P7	2.6	22	17
P8	2	12	10
P9	2.5	26	16
P10	3.4	19	19

TABLE II. EXPERIMENTAL RESULTS FOR PROPOSED - FOVSL

Polygon	Proposed technique (FoVSL)		
	Tour Length (in km)	Total number of turns	Execution Time (in s)
P1	1.3	10	6
P2	1.5	10	12
P3	1.4	11	15
P4	1.5	8	9
P5	2.4	18	12
P6	2.8	20	17
P7	2.4	20	14
P8	1.9	12	10
P9	2.4	23	14
P10	3.1	14	18

The comparison shows that the number of turns taken by the UAV has reduced drastically in the proposed method. Considering the sample polygons, a minimization of about at least 0.1km is obtained with respect to the total distance travelled by the UAV. The maximum difference is about 0.9km and the difference varies with respect to the number of sub-polygons obtained. While considering the time consumption for traversing the area, the minimization is evident for the proposed technique. For all the polygons taken, the time consumed by the proposed technique is minimal when compared with the existing technique and hence can be claimed to be more efficient than the existing algorithm.

V. CONCLUSIONS

Path planning of UAVs has become one of the most researched topics as they are now been deployed in various applications. This paper finds an optimal path over a given geographic area in the presence of static obstacles. Initially, the geographic area is divided into convex sub-polygons in which all the polygons need to be visited by the UAV. Once the decomposition is over, the proposed FoVSL technique is used over the decomposed polygons to obtain the optimal path in presence of static obstacles. Comparisons are done with the existing Sweep-Line method and results are tabulated for ten sample polygons. The comparisons prove that the proposed algorithm outperforms the existing

technique in terms of the total distance traveled, the number of turns taken by the UAV, and the total time consumed for the entire traversal. Future work could also include path planning considering dynamic obstacles and also the direction of the wind.

REFERENCES

- [1] Al-Kaff, A., Martin, D., Garcia, F., de la Escalera, A., & Armingol, J. M. (2018). Survey of computer vision algorithms and applications for unmanned aerial vehicles. *Expert Systems with Applications*, 92, 447-463.
- [2] Al-Kaff, A., Martin, D., Garcia, F., de la Escalera, A., & Armingol, J. M. (2018). Survey of computer vision algorithms and applications for unmanned aerial vehicles. *Expert Systems with Applications*, 92, 447-463.
- [3] Outay, F., Mengash, H. A., & Adnan, M. (2020). Applications of unmanned aerial vehicle (UAV) in road safety, traffic, and highway infrastructure management: Recent advances and challenges. *Transportation research part A: policy and practice*, 141, 116-12.
- [4] Zhang, H., Xin, B., Dou, L. H., Chen, J., & Hirota, K. (2020). A review of cooperative path planning of an unmanned aerial vehicle group. *Frontiers of Information Technology & Electronic Engineering*, 21(12), 1671-1694.
- [5] Priyadarsini, P. L. K. (2021, January). Area Partitioning by Intelligent UAVs for effective path planning using Evolutionary algorithms. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.
- [6] Hu, Y., Zhang, Z., Yao, Y., Huyan, X., Zhou, X., and Lee, W. S. (2021). A bidirectional graph neural network for traveling salesman problems on arbitrary symmetric graphs, *Engineering Applications of Artificial Intelligence*, 97(1):104-121.
- [7] Balampanis, Fotios, Maza, I., and Ollero, A. (2017). Area partition for coastal regions with multiple UAS, *Journal of Intelligent & Robotic Systems*, 88(2):751-766.
- [8] Karbowska-Chilinska, J., Koszelew, J., Os-trowski, K., Kuczynski, P., Kulbiej, E., and Wolejsza, P. (2019). Beam search algorithm for ship anti-collision trajectory planning, *Sensors*, 19(24):5338-5351.
- [9] Cabreira, T. M., Di Franco, C., Ferreira, P. R., and Buttazzo, G. C. (2018). Energy-aware spiral coverage path planning for UAV photogrammetric applications, *IEEE Robotics and Automation Letters*, 3(4):3662-3668.
- [10] Coombes, M., Fletcher, T., Chen, W.-H., and Liu, C. (2018). Optimal polygon decomposition for UAV survey coverage path planning in wind, *Sensors*, 18(7):2132-2160.
- [11] Reinelt, G. (1991). TSPLib—A Traveling Salesman Problem Library, *ORSA Journal on Computing*, 3(4):376-384.
- [12] Kapoutsis, A. C., Chatzichristofis, S. A., and Kos-matopoulos, E. B. (2017). DARP: Divide Areas Algorithm for optimal multi-robot coverage path planning, *Journal of Intelligent & Robotic Systems*, 86(3):663-680.
- [13] Nielsen, L. D., Sung, I., and Nielsen, P. (2019). Convex decomposition for a coverage path planning for autonomous vehicles: Interior extension of edges, *Sensors*, 19(19):4165-4182.
- [14] Azpúrua, H., Freitas, G. M., Macharet, D. G., and Cam-pos, M. F. (2018). Multi-robot coverage path planning using hexagonal segmentation for geophysical surveys, *Robotica*, 36(8):1144-1166.
- [15] Fevgas, G., Lagkas, T., Argyriou, V., & Sarigiannidis, P. (2022). Coverage Path Planning Methods Focusing on Energy Efficient and Cooperative Strategies for Unmanned Aerial Vehicles. *Sensors*, 22(3), 1235. Li, J., Xiong, Y., She, J., & Wu, M. (2020).
- [16] A path planning method for sweep coverage with multiple UAVs. *IEEE Internet of Things Journal*, 7(9), 8967-8978.
- [17] Akshya, J., & Priyadarsini, P. L. K. (2020). Graph-based path planning for intelligent UAVs in area coverage applications. *Journal of Intelligent & Fuzzy Systems*, 39(6), 8191-8203.

A Novel Multi Carrier PWM Based Harmonic Mitigation of Multi Level PV Inverter with Grid Integration

Satyanarayana.vanam

Asst.Professor

Department of EEE

Vaagdevi college of Engineering,

Warangal, Telangana, India

saisatyamvanam@gmail.com

Manne navya sree

M.Tech (PE), student

Department of EEE

Vaagdevi college of Engineering,

Warangal, Telangana, India

mennenavya123@gmail.com

ABSTRACT: Present paper tells about the enhanced PWM method for a five level transformer with fewer grids connected and standalone diode clamped multilevel inverter (DCMLI) for reducing the leakage current magnitude. Here we use 2 modified reference waves with improved PWM approach's for producing five level inverter output terminals by gaining the enhanced PWM approach uses single triangle wave. Here to decrease the total harmonic distortion (THD) in the grid current we develop the harmonics to multiple carrier frequencies. The inverter terminals are diminished when the harmonics are shifted to higher order according to the size of the required filter. To synthesize the five level ac output, just 24 power semiconductor switches are required and other six are operated at line frequency so that it diminishes the switching loss and also the system total cost to handle power imbalance among PV sources for every phase of 3-phase inverter to employ a new modulation strategy in its regulation. So the grid connected 5-level multi string inverter gives some interesting results using grid integration.

KEYWORDS: *PWM based multi level PV inverter, Total harmonic distortion (THD), Hybrid Bridge, PV sources, grid integrations*

Introduction:

A multi level inverter is made up by a collection of switching and DC source devices so that the output is stepped waveform by a amount of DC levels. In the last few decades, numerous topologies by a wide range of controller approaches have been established. Traditional multi-level inverters, such as the Clamping Hybrid bridge has certain disadvantages. As a

result, extensive study is underway to build Smart multilayer inverters by significant benefits. NPC employs a high amount of clamping diodes through varying voltages; in FC, switching efficiency is low due to the huge amount of capacitors, and the unstable the topology's fundamental flaw. However, as related to NPC and FC topologies, cascaded Hybrid bridge provides enhanced output THD and reduced distortion, however it needs a high amount of dc sources [1]-[5].

Aside since these 3 traditional topologies, a slew of new ones are introduced. A great number of topologies have recently been presented that are built on hybrid multistage and asymmetrical arrangements. Bidirectional switches are utilized in hybrid multistage configurations to improve the presentation of 1-phase and 3-phase MLI topologies [6]-[12]. Bidirectional switches, as opposed to unidirectional switches, can conduct current and tolerate voltage in both ways. Through the right controller method, bidirectional switches may increase the performance of multilevel inverters by decreasing the amount of lowering the bearing voltage, and attaining the wanted output voltage at higher levels [6]-[8]. dc supply is inadequate in asymmetrical designs. However, these arrangements are more dependable with fewer switches, sources, and capacitors, as well as a smaller installation space and lower cost [11]-[14]. [6] Proposes a novel 27 level MLI topology that optimizes the amount of output voltage levels while having a THD. [7] Proposes a novel (MLDCL) and a bridge inverter. A diode

clamped phase leg, a flying-capacitor phase leg, or cascaded half-bridge cells by own DC sources can all be used to realize MLDCL. Despite the switches' greater overall VA rating, this design requires 4 heavy assessment switches for the output side Hybrid bridge. The literature [8] suggests a 19-level MLI topology depend on the series connection of sub multilevel inverters.

PV Configuration

A novel DC voltage sources, and standing voltage on the switches has been proposed by 2 innovative techniques that are formulated for the magnitude of DC voltage sources to project in this study [9], however present topology also necessitates bi-directional switches. [10] Proposes a novel cascaded topology by a smaller amount of passive devices and shielded gate driving circuits.

This suggested architecture necessitates the use of four high-capacity bidirectional and Hybrid Bridge switches on the output side. This research employs a novel multilayer inverter design that generates a low THD output voltage by fewer power switches and dc sources. Innovative switching mechanisms for separate switches in the inverter reduce THD levels for greater output voltage.

The inverter, and the inverter is then interfaced with the Grid to meet Grid circumstances [13],[14]. Lastly, the model outcomes of a 5-level multilevel inverter, as well as Grid voltage and current forms subsequently interface, are included in the study. A classic PV cell is made up of n PV cells connected together.

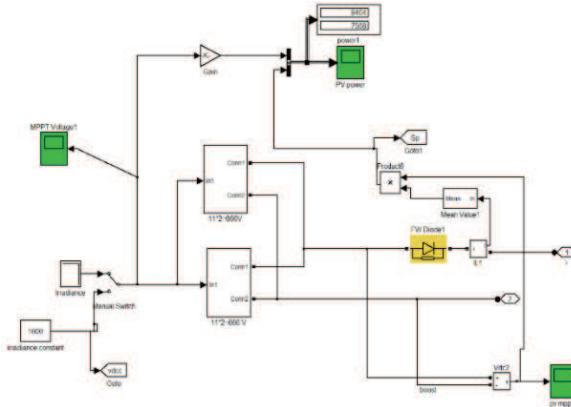


Fig.1 A classic PV module prototype with MPPT (P&O)

The Projected Multi Carrier PWM PV Topology

The basic arrangement of a diode clamped Multilevel inverter is revealed in Figure 2. The multi-level converter architecture contains a sequence of n diode clamped converters. The converter could create an output voltage v_H by n level is considered cells by the similar dc-link voltage. This high-quality voltage waveform allows for a decrease in harmonics in the filter inductor current, as well as a reduction in input filtering effort.

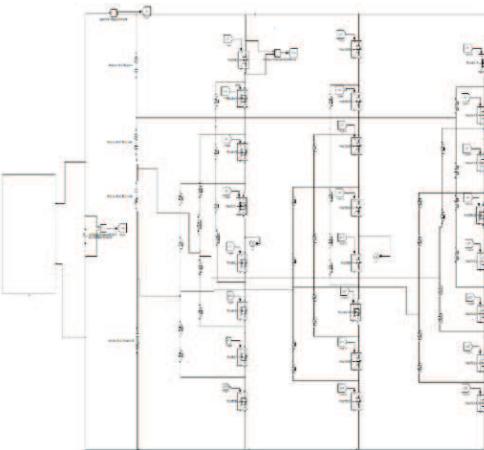


Figure 2. The PV based diode clamped multi-level converter topology

Many PWM strategies for multi-level inverters have been presented over the years [12],[13], with approaches depending on sine-triangular assessment playing a large role, such as carrier phase-shifted PWM patterns (PSPWM), carrier personality PWM patterns (CD-PWM), and composite carrier character PWM patterns (CCDPWM). The yield phase voltage of an inverter below PS-PWM controller is expressed as follows: [13]

$$v_a = nm_a v_{in} \sin \omega_s t + \frac{2nv_{in}}{\pi} \sum_{m=N,2N,\dots}^{\infty} \sum_{n=1,2,\dots}^{\infty} \frac{J_n(m m_a \pi)}{m} \cos m \pi \sin [(2m m_f + n) \omega_s t]$$

(1)

In this equation (1), n is the sum of Hybrid bridge components in the inverter, v_{in} denotes the amplitude of every components of DC source, and w_s denotes the sinusoidal modulation signal's radian frequency. When m_f is an integer, the yield voltages do not comprise even harmonics, as shown by equation (1). When m_f is not an integer, though, the voltage may contain even harmonics. The assessment of multi-level inverter is given as

$$m_f = \frac{f_{car}}{f_{ref}} \quad (2)$$

$$m_a = \frac{A_{ref}}{(N-1)A_{car}} \quad (3)$$

Where A_{car} is the carrier signals amplitude; f_{car} is the carrier signal's frequency; A_{ref} Left and right legs make up the Hybrid bridge topology. The Hybrid bridge output voltage waveform is twice as fast as every leg's

The phase difference among the upper and lower Hybrid bridges is 90°. It's possible that the switching frequency of every Hybrid bridge is doubled. In accumulation, as illustrated in fig. 3, the output voltage of 2 Hybrid bridge topologies is a 5 level waveform. ($m_f=35$, $m_a=0.8$) $m_f=35$, $m_a=0.8m_f=35$,

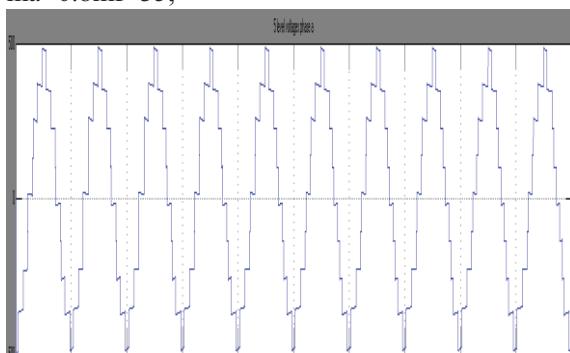


Fig. 3. The 5-level output voltage of inverter

Level Shifting PWM

The carrier of the components in all level-shifted PWM techniques wherever the carrier signals frequency is contrariwise proportionate to the

device's switching time On the other hand, the reference voltage might range from $-MV_{de}$ to MV_{de} .

The carriers are adjusted vertically to shield the whole voltage range, with the first module's carrier covering the range by 0 to V_{dc} and the 2nd covering the assortment by V_{de} to $2V_{de}$. The final element deals with voltages ranging by $(M-1)V_{de}$ to MV_{de} . Here are 3 different types of level shifting modulation methods:

- Phase Opposition Disposition
- Alternative Phase Opposition Personality
- Phase Disposition

The POD and APOD Entire carriers in the through total crews. When compared to the other disposition approach, Although level shifted multicarrier modulation provides superior harmonic attenuation, it does so at the expense of unequal device conditions. Fig 4: Carrier arrangements for POD 7 gating pulses of 5 level inverter

Simulation Modelling & Analysis

Figure 5 depicts a simulation figure of the projected grid-connected PV system element. It contains of a PV structure and a novel multi-level inverter for grid integration. Fig.5 Grid-connected photovoltaic system with five MLI levels. The PV cell converts solar energy directly hooked on direct current power.

The suggested inverter converts the voltage acquired by the PV into ac. lastly; the suggested inverter is associated to the electricity grid while meeting grid necessities like grid voltage phase angle, frequency, and amplitude.

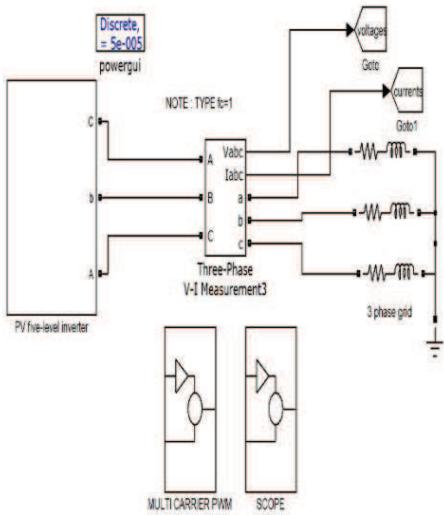


Fig.5 simulation model of multi carrier PWM based multi level PV inverter

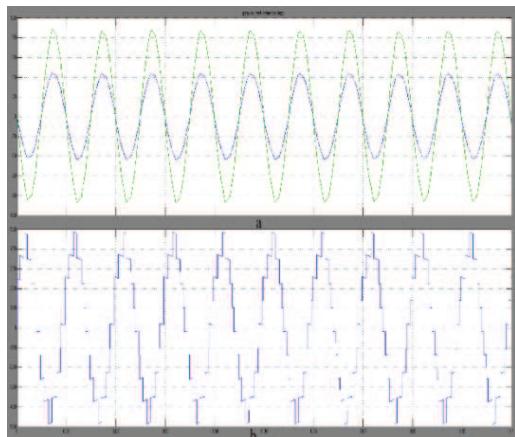


Fig 6 AC Current. (a) Step alteration in the AC currents: top $i_a(t)$, $v_a s(t)$, bottom $v_{ab}(t)$.

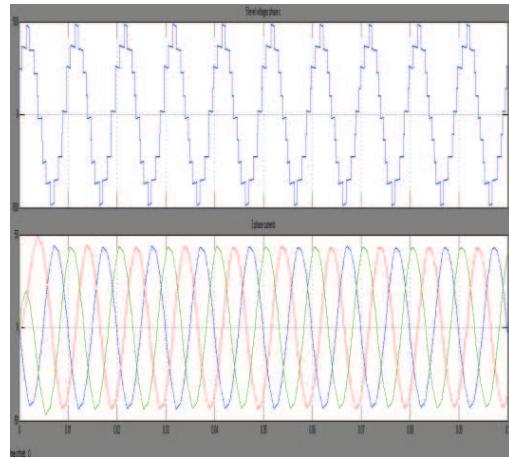


Fig.7 Key waveforms connected to the current control; (a) $i_{abc}(t)$. (b) $v_{ab}(t)$,

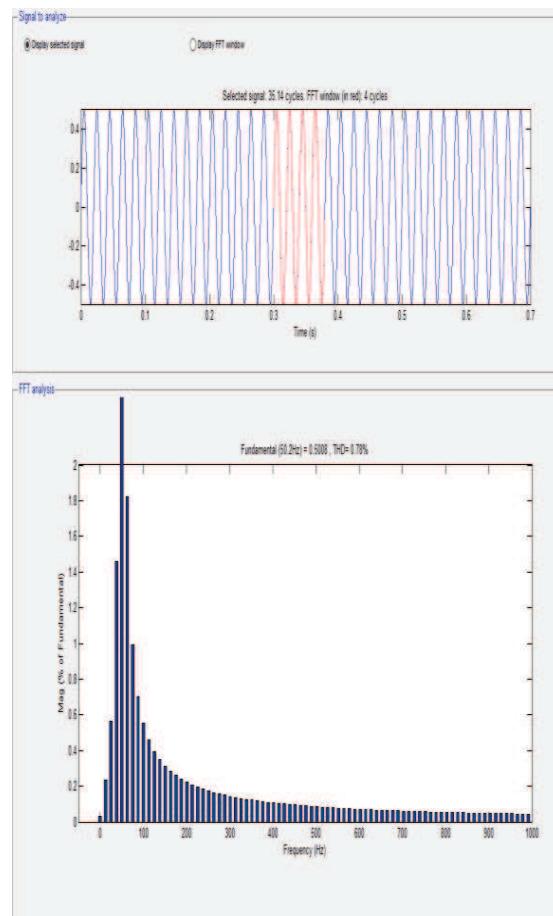


Fig .8 FFT analysis of inverter voltage (THD 0.78%)

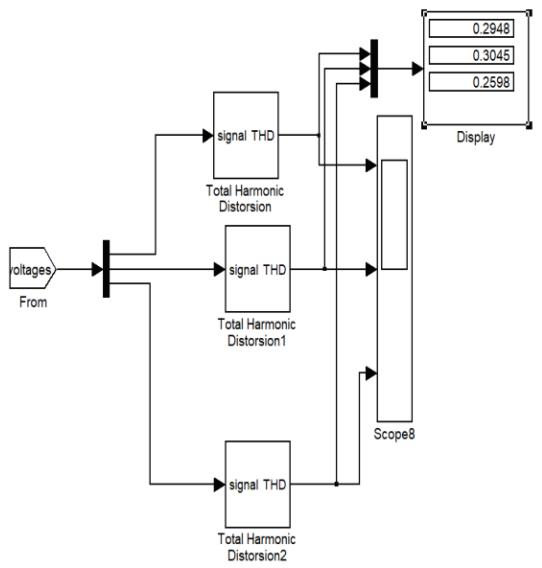


Fig.8 THD total harmonic distortion
0.29%,0.30%,0.25%

Conclusion

In this research, a novel architecture by a PV PCS based on a 5-level diode lamped topology by great efficacy is given. The projected topology diminishes current THD, filter size, and system efficacy significantly. The system efficacy in Europe is equal to 99.2 percent, with a peak efficacy of 99.6 percent. With the addition of the MPPT algorithm, the suggested topology's control algorithm manages the output current by unity power factor as well as the upper and lower PV input voltages. The investigational consequences suggest that the projected architecture performs well. It may also be utilized for added high-capacity renewable energyplications.

References:

- [1] Jong-pil Lee, Byung-duk Min, Tae-jin Kim, Dong-wook Yoo and Ji-yoon Yoo,"A Novel Topology for Photovoltaic DC/DC Full Bridge Converter with Flat Efficiency under Wide PV Module Voltage and Load Range." IEEE Trans. Industrial Electronics., Vol. 55, No. 7, pp. 2655-2663, July 2008.
- [2] Byung-duk Min, Jong-pil Lee, Jong-hyun Kim, Tae-jin Kim, Dong-wook Yoo and Eui-ho Song,"A New Topology With High Efficiency Throughout All Load Range for Photovoltaic PCS." IEEE Trans. Industrial Electronics., Vol. 56, No.11, Nov 2009.
- [3] Jong-pil Lee, Byung-duk Min, Tae-jin Kim, Dong-wook Yoo and Ji-yoon Yoo,"Design and Control of Novel Topology for Photovoltaic DC/DC Converter with High Efficiency under Wide Load Range." Journal of Power Electronics., Vol. 9, No.2, Mar 2009.
- [4] Jong-pil Lee, Byung-duk Min, Tae-jin Kim, Dong-wook Yoo and Ji-yoon Yoo,"Input-Series-Output-Parallel Connected DC/DC Converter for a Photovoltaic PCS with High Efficiency under a Wide Load Range." Journal of Power Electronics., Vol. 10, No.1, Jan 2009.
- [5] R. M. Santos Filho, P. F. Seixas, P. C. Cortizo, L. A. B. Torres, and A. F. Souza, "Comparison of Three SinglePhase PLL Algorithms for UPS Applications," IEEE Trans. on Industrial Electronics, Vol. 55, Issue 8, pp. 2923~2932, Aug. 2008.
- [6] Villava. M.G, Gazoli. J.R, Filho. E.R, "Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays," Trans. on Power Electronics, Vol. 24, No. 5, pp. 1198~1208, May. 2009.
- [7] Weidong Xiao, Ozog, N., and Dunford, W.G., "Topology Study of Photovoltaic Interface for Maximum Power Point Tracking," IEEE Trans. on Industrial Electronics, Vol. 54, No. 3, pp. 1696~1704, June 2007.
- [8] Rong-Jong Wai, Wen-Hung Wang, Chung-You Lin, "High-Performance Stand-Alone Photovoltaic Generation System," IEEE Trans. on Industrial Electronics., vol. 55, no. 1, pp. 240-250, Jan 2008.
- [9] S. Alepuz, S. Busquets-Monge, J. Bordonau, J. Gago, D. Gonzalez, J. Balcells, "Interfacing Renewable Energy Sources to the Utility Grid Using a Three-Level Inverter," IEEE Trans. on Industrial Electronics., vol. 53, no. 5, pp. 1504-1511, Oct 2006.

- [10] R. Leyva, C. Alonso, I. Queinnec, A. Cid-Pastor, D. Lagrange, L. Martinez-Salamero, "MPPT of photovoltaic systems using extremum-seeking control," *IEEE Trans. Aerospace and Electronic Systems.*, Vol. 42, Issue 1, pp. 249~258, Jan. 2006.
- [11] Villanueva, E., Correa, P.Rodriguez, J.Pacas, M., "Control of a Single-Phase Cascaded Hybrid bridge Multilevel Inverter for Grid-Connected Photovoltaic Systems," *IEEE Trans. Industrial Electronics.*, Vol. 56, Issue 11, pp.4399~4406, 2009
- [12] Keivani, H., et al. "Novel Multi-Carrier PWM Method for a Three-Phase Cascaded Hybrid bridge Multi-Level Inverter." in Universities Power Engineering Conference, 2006. UPEC '06. Proceedings of the 41st International. 2006.
- [13] Tengfei, W. and Z. Yongqiang. "Analysis and comparison of multicarrier PWM schemes applied in Hybrid bridge cascaded multi-level inverters." in Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on.
- [14] Dell'Aquila, A., et al., "Overview of PI-Based Solutions for the Control of DC Buses of a Single-Phase Hybrid bridge Multilevel Active Rectifier." *IEEE Trans., Industry Applications*, 2008.
- [15] Nabae A, Takahashi I, Akagim H. A new neutral point clamped PWM inverter. *IEEE Trans Ind Appl* 1981;IA-17(5):518–23.
- [16] Krishna Kumar Gupta, Shailendra Jain A multilevel Voltage Source Inverter (VSI) to maximize the number of levels in output waveform, *Electrical Power and Energy Systems* 44 (2013) 25–36.
- [17] S. Gui-Jia, Multilevel dc-link inverter, *IEEE Trans. Ind. Appl.* 41 (3) (2005), 848–845
- .
- [18] E brahimBabaei, Ali Dehqan, Mehran Sabahi A new topology for multilevel inverter considering its optimal structures, *Electric Power Systems Research* 103 (2013) 145– 156.
- [19] E. Babaei, Optimal topologies for cascaded sub-multilevel converters, *J. Power Electron.* 10 (3) (2010) 251–261.
- [20] E. Babaei, A cascade multilevel converter topology with reduced number of switches, *IEEE Trans. Power Electron.* 23 (6) (2008) 2657–2664 .
- [21] J. Dixon, L. Morán, High-level multistep inverter optimization using a minimum number of power transistors, *IEEE Trans. Power Electron.* 21 (2) (2006)330–337.

Power Quality Improvement in Grid System with PV Based SVPWM-DVR for Sag & Swell Mitigation with a Novel ANN Controller

Satyaranayana.vanam
Asst. Professor
Department of EEE
Vaagdevi college of Engineering,
Warangal, Telangana, India
saisatyamvanam@gmail.com

B.SHIVA PRASAD BHARATH
M.Tech (PSCA), student
Department of EEE
Vaagdevi college of Engineering,
Warangal. Telangana, India
banavathshiva44@gmail.com

Abstract: This research proposes the since of semiconductor devices, the quality of electricity delivered to consumers is deteriorating. Here maximum prevalent power quality issues in the distribution system are voltage sag, voltage swell, and harmonics. The Dynamic Voltage Restorer (DVR) is a specialized power device that might be the finest answer for improving the efficiency and effectiveness of the power system. DVR stores the energy in a storage element that uses alternate energy from a Solar PV cell. To balance solar power, a P&O method depended on Maximum Power Point Tracking (MPPT) is utilized so the sensitive power is generated by the Voltage Source Inverter (VSI), which must be corrected using space vector PWM (SVPWM) and the feedback control loop to improve the injection of reactive power into the line. Present research is about a photovoltaic source depended DVR system that combines nonlinear inverter controlling methods like ANN and hysteresis voltage method for restoring voltage sag, swell and also the fault circumstances that occurs in structure by complete revision utilizing MATLAB electrical simulation software.

Keywords: voltage sag and swell, THD, DVR, PV, ANN, SVPWM and APPT based P & O algorithm

Introduction:

Renewable energy sources like solar and wind are fast growing as an alternate energy source to create power due to a scarcity of fossil energy sources and rising alertness of environmental implications. PV generates DC electricity, which necessitates the use of an inverter formerly it can be used and associated to the distribution network. Although PV may deliver power to the grid, it has a flaw: it generates harmonics as a result of the presence of a voltage source inverter (VSI) on the active filter series, lowering power quality. As well as sensitive loads become more prevalent, the number of power quality issues in the distribution method has improved dramatically. Sag, swell, and short-circuit are the most serious and common grid voltage disruptions. The sag is defined as a drop in rms voltage of 10% to 90% that lasts since 1/2 cycle to 1 min. The swell is a short-term rise in source rms voltage that ranges as of 1.1 pu to 1.8 p.u over insignificant source voltage. DVR is a gadget that efficiently and thoroughly mitigates sag/swell fault. [1] A study of sag compensation utilizing DVR and the UVTG technique was conducted. The DVR can pay off for both

balanced and unbalanced sag and inject a preferred voltage constituent alongside source voltage to quickly repair a variety of disruptive variances and retain load voltage persistent and balanced at nominal value.

Though, due to the existence of sensitive voltage devices on the load side, that research did not address voltage swell and harmonic abatement. A comparison of compensation strategies, such as phase recompense, pre dip recompense (also known as recompense of significant voltage variance), and intelligent phase recompense is conducted [2].

A low-rating DVR, according to IEEE 519, could recompense for both sag and swell to decrease voltage THD on the load bus. [4] used Sinusoidal Control PWM (SPWM) and Space Vector PWM to study the performance of DVRs in terms of sag and load voltage harmonics (SVPWM). The Synchronous Reference Frame (SRF) technique is utilized to identify sag voltage and produce modulation signal.

SVPWM has a lower load voltage THD than SPWM, according to the research. The presentation of SVPWM and SPWM by load voltage swell, however, is not addressed in present paper. The combination of a grid-connected PV structure and a self-supported DVR is projected [5]. The structure is known as the "six-port converter," and it is made up of nine semiconductor switches, as opposed to the prior 12 semiconductor switches. Depending on grid conditions and PV generation power, the designs can operate in a variety of modes. Normal grid approach, fault approach, sag approach, and non-active PV approach are all research modes.

Problem Formulation

In radial systems, the voltage divider model is widely used to compute the magnitude of voltage sag/swell at the point of common coupling (PCC) (that is the supreme frequent in industrial distribution networks). Conferring to this approach, power quality problems are

characterized as presented in figure 1, here the voltage magnitude at the PCC is specified by:

$$V_{\text{sag swell}} = Z_s + Z_f \quad (1)$$

The source impedance, which includes the transformer impedance, is Z_s , and the impedance among the PCC and the fault, which includes the fault and line impedances, is Z_f .

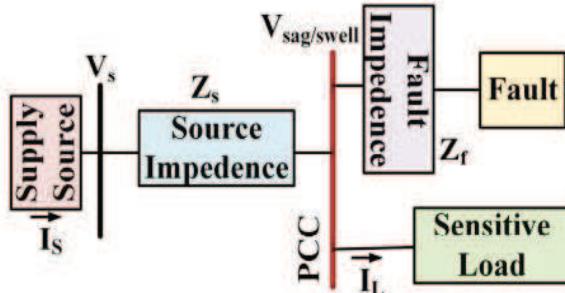


Fig 1. To show power quality issues, a block diagram of a power system is used.

Dynamic Voltage Restorers (DVR) are complicated fixed devices that supply the essential voltage to reestablish the amplitude to a steady zone throughout a voltage sag. To put it another way, the device inserts power into the structure to reestablish the voltage to the load's required level. To inject power, a switching mechanism is employed in aggregation through a transformer associated in series by the load.

DVRs can be divided into two categories: those by energy storage and those deprived of devices to reestablish the voltage waveform without energy storage by illustrating the needed current as of the source. Here alternate type stores energy to adjust for voltage sag.

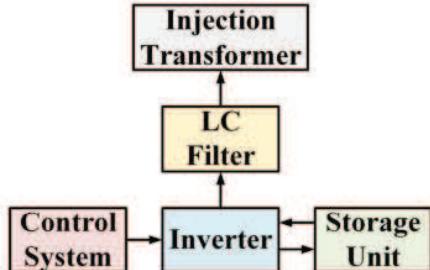


Fig 2. Schematic of Dynamic Voltage Restorer (DVR).

DVRs can be divided into two categories: those with energy storage and those without. Devices reestablish the voltage waveform without energy storage by illustrating the needed current as of the source. A DVR with storage differs from a UPS in that it gives the segment of the waveform to reduce in amplitude owed to voltage sag, rather than the complete waveform. Here a UPS is a power supply, however a DVR is only a compensator for

any power interruption from the source. This is depicted in Figure 2.

Vaccinate /supporter transformer, a vocal filter, a voltage source converter (VSC), and a control system make up a basic DVR. DVR systems are extremely effective and quick to respond. None of the inherent storage difficulties are important in the event of structures deprived of storing. Another important feature of DVR systems is that, in addition to voltage sag prevention, they may be utilized for harmonic extenuation, fault current preventive, power factor alteration, and transient decrease.

Dynamic voltage restorers use three voltage sag compensation strategies to handle power eminence issues: pre-sag recompense, in-phase recompense, and phase progressive recompense methodology. In the curve illustrated in figure 3, $V_{\text{pre-sag}}$ and V_{sag} represent voltages at the point of common coupling (PCC) formerly and throughout the sag, correspondingly. DVR injects a voltage to recompense for pre-sag, which is given as:

$$|V_{\text{inj}}| = |V_{\text{pre-sag}}| - |V_{\text{sag}}| \quad (2)$$

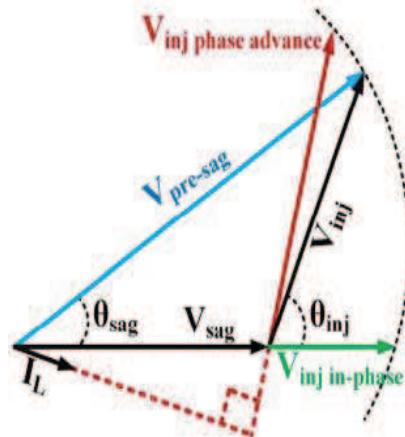


Fig. 3. Sag recompense strategies are depicted as a phase diagram. The compensation voltages for the three different procedures are represented by the blue, red, and green vectors.

$$\theta_{\text{inj}} = \tan^{-1} \left(\frac{V_{\text{pre-sag}} \sin(\theta_{\text{pre-sag}})}{V_{\text{pre-sag}} \cos(\theta_{\text{pre-sag}}) - V_{\text{sag}} \cos(\theta_{\text{sag}})} \right) \quad (3)$$

The voltage inserted by DVR for in-phase recompense method is specified by:

$$V_{\text{DVR}} = V_{\text{inj}} \quad (4)$$

$$|V_{\text{inj}}| = |V_{\text{pre-sag}}| - |V_{\text{sag}}| \quad (5)$$

$$\angle V_{\text{inj}} = \theta_{\text{inj}} = \theta_{\text{sag}} \quad (6)$$

The fundamental idea behind the phase advance recompense technique is to exclusively use reactive power injection to achieve the desired compensation. The inserted voltage and load current are 90 degrees apart in this manner.

Projected methodology

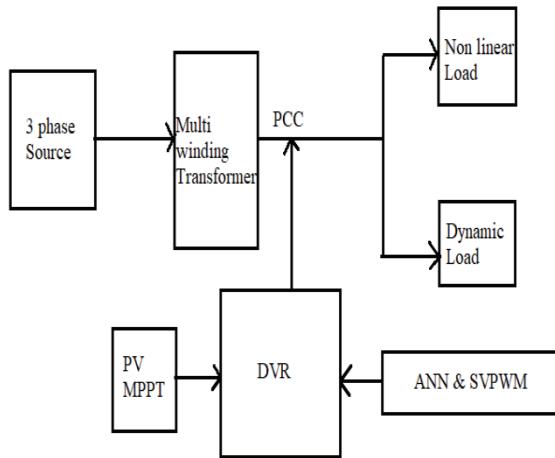


Fig. 4 ANN based PV-DVR Proposed schematic

A photovoltaic source dynamic voltage restorer in a 3-phase structure employing Hysteresis voltage and ANN controllers is exposed in the illustration overhead. When a voltage sag, swell, or fault develops in the structure, the suggested controller approaches detect it and send control signals to the inverter, causing it to provide voltage of the essential magnitude, thereby minimizing voltage sag, swell, and fault conditions. A renewable energy source, such as a solar module, provides the active power essential for the inverter to produce compensating voltage.

A. PV module

Photovoltaic cells convert visible light into direct current, and solar panels or arrays could be created by attaching PV cells in series or parallel, depending on the application. The main benefit of photovoltaic is that it is environmentally beneficial and that solar energy is limitless. When installed, the PV component offers energy at a low cost and requires little maintenance. The PV model is designed using a variable irradiance in this study.

Before utilizing MPPT, a PV module's output power is 30kW, while its voltage output is 800V. Incremental conductance is one of the MPPT algorithms that are utilized to trap maximum power from PV modules by lowering the voltage to the appropriate level. Then the PV segment is connected to the DVR's three-phase inverter to deliver real power for voltage correction throughout sag and swell as well as fault scenarios.

B. Hysteresis voltage controller

In the event of voltage sag, swell, or failure, hysteresis voltage control uses feedback to alter the inverter's output and provide compensating voltage. The Hysteresis voltage control method works by creating an error signal by relating the actual voltage (restrained) and reference voltage signals, then using a Hysteresis comparator to trigger inverter switching pulses.

A voltage sag of magnitude 0.8 p.u for the period of 0.1 to 0.15 seconds, a voltage swell of magnitude 1.2 p.u for the time range of 0.15 to 2 seconds, and an superficially applied 3 phase fault for the period of 0.25 to 0.3 seconds are used to calculate the presentation of the photovoltaic source built dynamic voltage restorer in a structure by projected Hysteresis voltage control method.

C. ANN organizer

A neural network (ANN) is a computer control mechanism that consists of numerous types of small pieces functioning in parallel, related to how the human biological nervous system detects models [3]. This neural network obtains the finest possible outcome by responding to changes in input deprived of changing the output circumstances. Traditionally, neural networks have been taught to follow a given input to a specific intended output. The exogenous input in this suggested exertion is a nonlinear auto relapsing exogenous prototype that ties the current assessment of a time series to its previous assessment, as well as the current and preceding values of the lashing sequences. The network receives 7002 examples of three-phase voltage inputs and targets (21003 inputs and 21003 targets), which are unevenly divided into three sectors.

Typically, 70% of the network is utilized for exercise, 15% for authenticating the system for generalization, and 15% for challenging the network. Levenberg-Marquardt is a exercise procedure that requires a lot of memory yet takes a short amount of time. That simplification stops improving, the preparation comes to a stop, that is signaled by a rise in the mean square error for the validation samples. As a result, the mean square error should be reduced. The variance among the average values derived as of the output and the goal is known as the mean square error, and the R stands for regression,

which assesses the correlation between the output and the target. R value of 1 specifies a close link, though R value of 0 specifies an irregular relationship. The neural network has been effectively trained to predict the required output if the attained mean square error is 0.004436 and R is 0.99 respectively.

Artificial Neural Network

A mathematical or computer prototype depending on the erection and purpose of biological neural networks is recognized as artificial neural networks (ANN). These networks can rely on a huge amount of unknown inputs to estimate or approximate functionality. A hypothesis is formed, and the gradient output will be calculated using it. The following is an explanation of the computed network's output: The output of the concealed layer's node is as follows:

$$y_j = f(\sum_i w_{ji}x_i - \theta_j) - f(net_j) \quad (7)$$

Where,

$$net_j = \sum_i w_{ji}x_i - \theta_j \quad (8)$$

Computational output of the output node:

$$z_i = f(\sum_i v_y y_j - \theta_i) = f(net_i) \quad (9)$$

Where,

$$net_i = \sum_i v_y y_j - \theta_i \quad (10)$$

Error of the output node:

$$E = \frac{1}{2} \sum_i (t_i - z_i)^2 = \frac{1}{2} \sum_i (t_i - f(\sum_i v_y y_j - \theta_i))^2 \quad (11)$$

Hypothesis:

$$h_\theta(x) = \theta^T x = \sum_{i=0}^n \theta_i x_i \quad (12)$$

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (13)$$

Where

I x = input node,
j y = hidden layer node,
l z = output layer node,
w_{ij} = weight value of network between input and hidden layer node,
v_{iy} = weight value of network between hidden layer and output layer nodes,
t_i = expected value of the output node = learning rate,
m = total sample Figure 5 depicts the

ANN's training technique.

Even if the topology is not the same as that used in this study, the preparation may be enhanced to provide a proximate perfect outcome in all scenarios, straight if the topology is not the same as that used in this study. The training time can also be changed to account for the trade-offs among accurateness and the amount of nodes in the ANN; as the amount of layers and nodes in the ANN drops, so does the training time, but so does the accuracy. Fig.5. The ANN controller for DVR goes through a training procedure.

Simulation Consequences and also Deliberations

A 3-phase structure with a solar source and a dynamic voltage restorer, as well as a nonlinear load, is constructed to test the projected control strategies. The simulation experimentation was done in 3 phases by MATLAB/Simulink, including:

stage 1: With PR controller,

Stage 2: With ANN controller sag compensation,

Stage 3: With Artificial neural network swell compensation.

Stage 1: System with PR control.

Pretend circuit for a 3-phase structure that does not include a DVR or a PR controller within, a Sag, Swell, is formed with a 3-phase programmable source, and an externally applied three-phase fault is applied to the system. The voltage at the load side is altered as a consequence of these situations.

Figure 6 depicts a inaccurate voltage waveform at the source and load because of sag, swell, and fault circumstances. The load voltage THD for Sag, Swell, and Fault circumstances, respectively, is 62.42 percent, 62.37 percent, and 81.82 percent. As a result, a dynamic voltage restorer by SVPWM an ANN was offered as a way to improve the results, and it was tested further.

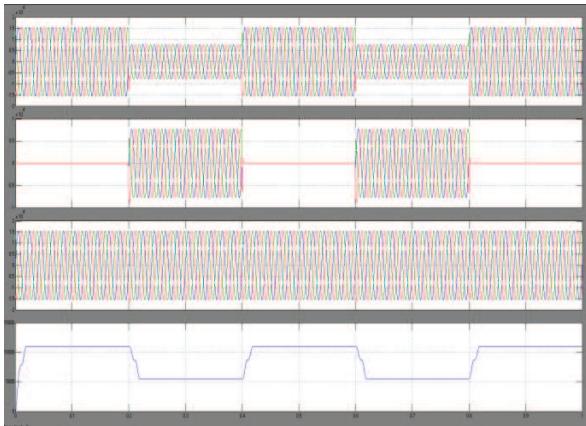


Fig. 6. Output waveform of inaccurate voltage at source and load side of the Structure with PR control

Stage 2: With ANN controller sag compensation,

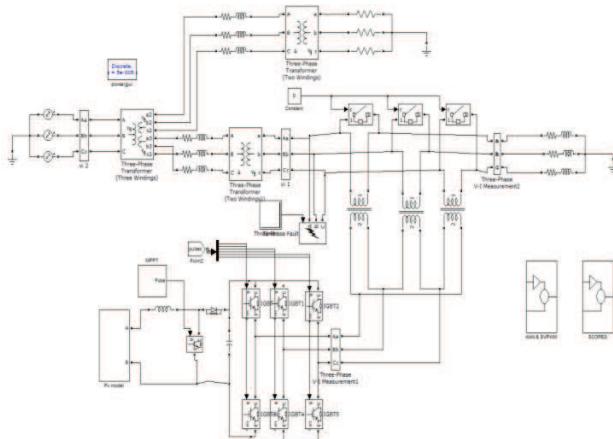


Fig. 7. PV based DVR with ANN control method voltage sag .

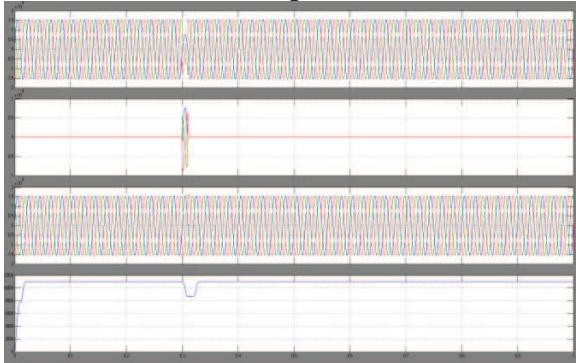


Fig. 8. Output waveform of inaccurate voltage at source and load side sag compensation of the Structure with SVPWM ANN control

Stage 3: With ANN swell compensation

A 3-phase structure by a dynamic voltage restorer and an ANN controller is shown in Figure 9. The voltage is

inaccurate due to sag and swell, as well as fault circumstances, which are adjusted using the projected prototype, which consists of a voltage source inverter, an ANN controller, and a solar component.

The results of the experiments show that THD on load side voltage for sag and swell circumstances is condensed to 2.83 percent and 3.44 percent, correspondingly (relate to Hysteresis controller). Figure 11 shows the inaccurate voltage waveform and the adjusted voltage waveform achieved by improving voltage quality with a dynamic voltage restorer and an ANN.

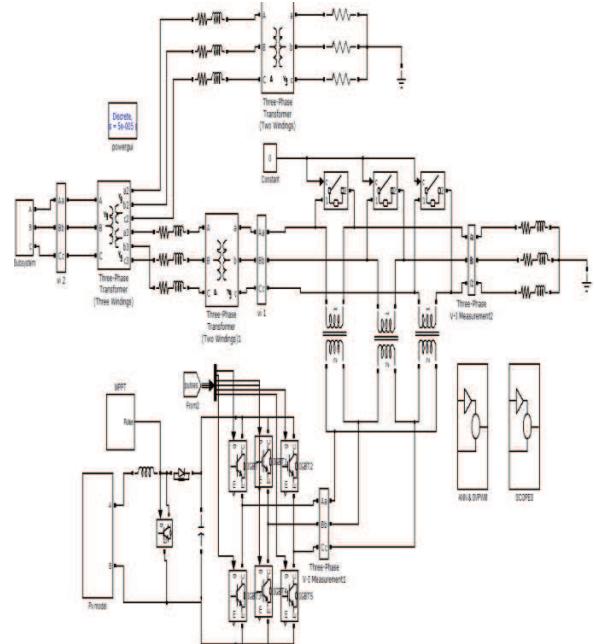


Fig. 9. PV based DVR with ANN control method voltage swell

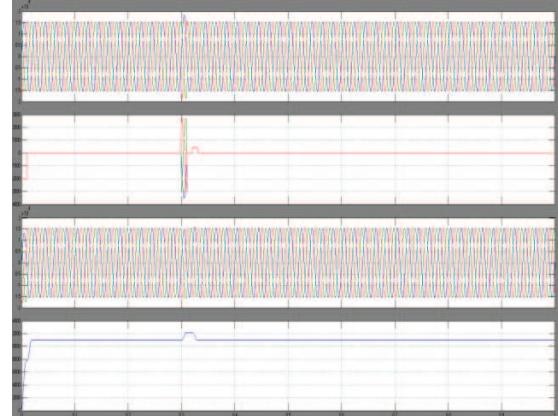


Fig. 10. Output waveform of inaccurate voltage at source and load side swell compensation of the Structure with SVPWM ANN control

Conclusion

Using an ANN and the Hysteresis voltage control method, this study creates and analyses a DVR with a photovoltaic source. To evaluate the suggested control options for the present three-phase system, a Sag, Swell, and 3-phase external fault are manufactured internally at the source. The Artificial neural network controller decreases voltage sag and swell well, according to simulation findings and THD analysis, whereas the Hysteresis voltage controller performs best for voltage adjustment in defective circumstances, thereby enhancing system voltage quality. When associating Hysteresis and ANN controller performance, the consequences openly show that ANN control method is the preeminent appropriate method for modifying Sag and Swell circumstances, and Hysteresis voltage control method is utilized to moderate Defective circumstances in the 3-phase system.

References

1. Miska Prasad,Ashok kumar kella,"Voltage & Currnet quality improvement by solar photovoltaic fed ZSI -DVR",pp. 434-441,ICSCC ,Elsevier,December 2017.
2. N. Srinivasa Raoa , Dr. A. Selwin Mich Priyadharsonb , Dr. J. Praveenc, "Simulation of Artificial Intelligent Controller based DVR for Power Quality Improvement",pp. 153-167,Procedia computer science,Elsevier,2015.
3. Md Samiul haque Sunny,Eklas Hossain, Mikal Ahmed, Fuad UnNoor, "Artifical Neural Network based Dynamic Voltage Restorer for Improvement of Power quality",pp. 5565-5572,IEEE,2018.
4. Nizamulmuluk,Agung Warsito,Juningtyastuti,Iwan setiawan,"Voltage sag mitigation due to short circuit current using dynamic voltage restorer based on hysteresis controller",ICITACEE, pp. 108-112,2017
5. Paramjit Kaur, santoshi Gupta "Mitigation Technique for Voltage Sag & Swell By Using Dynamic Voltage Restorer",IJREICE,vol 2,issue 1,pp 714-718, Jan 2014
6. Abrarkhan I.pathan,PROF.S.S Vanamane, "A Control Technique for Instant Mitigation of Voltage Sag/Swell by Dynamic Voltage Restorer",pp.2231-4407.
7. Thomas M. Bloming, P.E. and Daniel J. Carnovale, P.E., "Application of IEEE Standar 519-1992 Harmonic Limits", Presented at The 2005 IEEE IAS Pulp and Paper Industry Conference in Appleton, WI.
8. Arrilaga, Jos and Watson, Neville, "Power System Harmonics", Chichester: John Willey and Sons, 2003.
9. Tony Hoevenaar, P.Eng, Kurt LeDoux, P.E., Matt Colosino, 2003, "Interpreting IEEE Standard 519 and Meeting Its Harmonic Limit in VFD Application", Copyright Material IEEE Paper No. PCIC-2003- 15.May, 2003.
10. Sankaran C, "Power Quality", CRC Press LLC, 2002.
11. Amirullah and Agus Kiswantonono, "Power Quality Enhancement of Integration Photovoltaic Generator to Grid under Variable Solar Irradiance Level using MPPT-Fuzzy", International Journal of Electrical and Computer Engineering (IJECE), IAES Publisher, Vol. 6, No. 6, December 2016, pp.2629-2642.
12. C. Zhang, D. Zhao, J. Wang, and G. Chen, "A modified MPPT method with variable perturbation step for photovoltaic system," in Proc. Power Electronics and Motion Control Conference, pp. 2096-2099, 2009.
13. T. Noguchi, S. Togashi, and R. Nakamoto, "Short-current pulse-based maximum-power-point tracking method for multiple photovoltaic-and-converter module system," IEEE Trans. Ind. Electron., vol. 49, no. 1, pp. 217-223, Feb. 2002.
14. N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimizing sampling rate of P&O MPPT technique," in Proc. IEEE PESC, pp. 1945- 1949, 2004.
15. W. Wu, N. Pongrattananukul, W. Qiu, K. Rustom, T. Kasparis, and I. Batarseh, "DSP-based multiple peak power tracking for expandable power system," in Proc. APEC, pp. 525-530, 2003.
16. T. Esram and P.L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," IEEE Trans. Energy Conversion, vol.22, no. 2, pp. 439-449, June 2007.
17. K.H. Hussein, I. Muta, T. Hoshino, and M. Osakada, "Maximum photovoltaic power tracking algorithm for rapidly changing atmospheric conditions," IEE Proceedings on Generation, Transmission and Distribution, vol. 142, no. 1, pp. 59-64, Jan 1995.
18. S. Yuvarajan and S. Xu, "Photo-voltaic power converter with a simple maximum-power point-tracker," in Proc. ISCS, pp. 399-402, 2003

Hybrid Energy Management System & Power Compensation Of PV & Wind Integrated Grid System With Fuzzy Based D-STATCOM

Satyaranayana.vanam
Asst.Professor
Department of EEE
Vaagdevi college of Engineering,
Warangal, Telangana, India
saisatyamvanam@gmail.com

Sreedhar.gurijala
M.Tech (PSCA), student
Department of EEE
Vaagdevi college of Engineering
Warangal.Telangana, India
Sridhargurijala7@gmail.com

Abstract: Present paper is about a hybrid PV (Photovoltaic) and wind integrated controllers for distributed static compensator (DSTATCOM) grid ties method by adaptive reweighted zero appealing (RZA) method depending on perturb and observe (P & O) by max power point tracking system at 3-phase for improving the power quality and also for supporting the 3-phase AC grid system which is used for supplying both associated and grid loads. Present algorithm is capable of working according to clock in round independent of all climatic conditions and it also performs the dual functionality like improving the power quality to work as DSTATCOM and similarly it relocates the power to load and grid which is acquired by PV array and wind PMSG. As the system is working as independent in all climatic conditions for improving power quality like DSTATCOM it is termed as smart because it performs both methods repeatedly by identifying PV and wind power, so the skilled of multi-directional power flow. In the condition of an unstable and non-linear load, the DSTATCOM successfully provides stable and sinusoidal source currents by unity power factor (UPF). Fuzzy logic is used to sustain DC connection voltage. The controller gain of a fuzzy logic controller is designed to alter in response to dynamic load disturbances.

KEYWORDS: PV (photo voltaic) and fuzzy logic array system, MPPT P & O, DSTATCOM, RZA algorithm, voltage sag and swell

INTRODUCTION: Renewable energy systems, for example solar PV (Photovoltaic) and wind (PMSG), have achieved widespread acceptability due to the fact that solar PV and wind energy are both free and easily accessible [4]. To participate PV-wind systems by the electrical grid, the single phase topology, that links openly the PV method and wind by the grid and use a

converter, and the double stage topology, that attaches the PV structure and PMSG system by a converter to the grid by a boost converter, is reported in many studies [5]. The single stage topology, on the other hand, is preferred since it eliminates the need for a DC-DC boost converter, lowering costs and increasing power quality. PV wind-DSTATCOM (PV- Distribution Static Compensator) is a converter with such a PV array and an active filter feature that converts DC voltage into AC and improves power quality by restricting harmonic distortions, compensating sensitive power, and balancing power in all 3 stages of the distribution system [6]. Furthermore, the voltage and current of PV and wind have a nonlinear connection. As a result, there is a certain voltage and current at which a PV array and wind PMSG may provide max power [7]. Adaptive perturb and observe [9], grey wolf optimization methodology [10], neural network [11], and hybrid fraction open circuit voltage-current sensor minimal technique [8] are some of the max power point tracking (MPPT) strategies that is documented. Rising energy demands, outages, and a high saturation of nonlinear loads combined by an aged system is posed major grid issues [1]. To address these issues, this paper proposes a smart PV-WIND-DSTATCOM system that functions as an active filter and is capable of nullifying the effect of nonlinear load currents, regulating the DC-link voltage, improving the grid's power factor, and allowing multi - directional power flow. The suggested system can operate in 2 methods: PV-WIND-DSTATCOM and DSTATCOM, and can repeatedly execute in 2 ways of PV and wind power sensing. Depending on forecasts, an optimal scheduling approach may be devised to reduce the use of fossil fuels in power generation [12].

The PV-WIND -DSTATCOM combines the benefits of DSTATCOM with the ability to reduce current quality issues while also transferring power to the load, lowering the group demand from traditional sources. If there is enough power, a multi - directional power flow can occur, with the power being sent mutual to the grid and to the load. Here solar PV and wind power are unavailable, the system shifts to DSTATCOM mode, and electricity is supplied to the burden via the grid and the converter system.

System Configuration

Figure 1 depicts the system under investigation. Figure 1 depicts a 3-phase, four-wire (3p4w) distribution method. DSTATCOM depending on DC-link VSI is presented and considered. DSTATCOM is associated to the PCC via R_f and L_f interface inductors. DSTATCOM is synchronized by a VSI topology with two DC link capacitors C_1 and C_2 with V_{dc1} and V_{dc2} voltages. Isolated gate bipolar transistors (IGBTs) are utilized to implement VSI. In order to offer constant balancing,

the PV array is also associated to the DC side of the higher capacitor via a boost converter. Using the P & O approach, the o/p of a PV array is traced by highest power point. Minor dc link capacitor is strengthened using PMSG depending wing turbine to generate hybrid renewable energy depended dc-link.

A switch mode rectifier with appropriate control is utilized to capture the max power from the wind. The dc-link voltage is intended to be modulated between 1.2 and 2 intervals the structure peak voltage, because the percent THD of remunerated source current is determined by lowest whenever the dc voltage is about equivalent to 1.6 times the peak of the ac system voltage. Both renewable sources have their own control systems for regulating their separate dc link capacitor voltages. Table I lists the structure parameters utilized in the simulation process, whereas Appendix I lists the PV array and wind measurements.

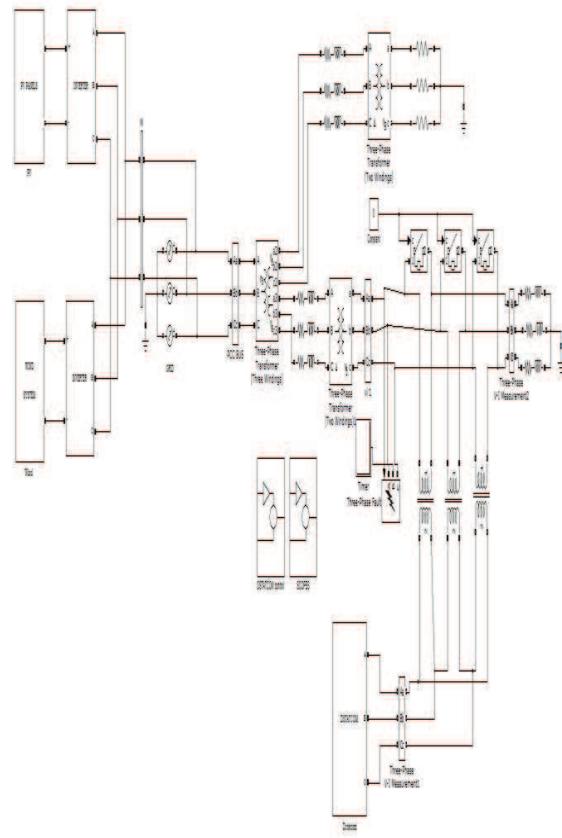


Fig.1 Proposed Fuzzy based DSTATCOM with hybrid power system

Modeling of PV Array

A. PV features an corresponding path of a concrete solo photovoltaic cell is exposed in figure Numerals of PV cell creating components exist

$$\left. \begin{aligned} i_{fa}^* &= \frac{V_{sa} + (V_{sb} - V_{sc})\beta}{V_{sa}^2 + V_{sb}^2 + V_{sc}^2} (P_{avg} + P_{loss}) \\ i_{fb}^* &= \frac{V_{sb} + (V_{sc} - V_{sa})\beta}{V_{sa}^2 + V_{sb}^2 + V_{sc}^2} (P_{avg} + P_{loss}) \\ i_{fc}^* &= \frac{V_{sc} + (V_{sa} - V_{sb})\beta}{V_{sa}^2 + V_{sb}^2 + V_{sc}^2} (P_{avg} + P_{loss}) \end{aligned} \right\} \quad (1)$$

TABLE I

SYSTEM PARAMETERS

System Parameters	Values (Ratings)
Source voltages	Balanced sinusoidal 230 V (rms), 50Hz
Unbalanced Linear Load	$R_{la}+jX_{la}$: 30+j22 Ω $R_{lb}+jX_{lb}$: 60+j31.4 Ω $R_{lc}+jX_{lc}$: 120+j125.6 Ω
Non-Linear Load	Three phase uncontrolled rectifier with R-L load of 150 Ω and 40 mH
2-level NPC inverter	$C_{dc1} = C_{dc2} = 3000 \mu F$, $V_{dc\ ref} = 750 V$ (each)
Interface inductor	$R_f = 0.01 \Omega$, $L_f = 40mH$
Hysteresis band	$HB = \pm 0.07 A$
Proportional Integral gains	$K_p = 0.01$, $K_i = 10$
Boost Converter	$L = 0.0191 mH$, $C = 3000 \mu F$ $F = 25kHz$

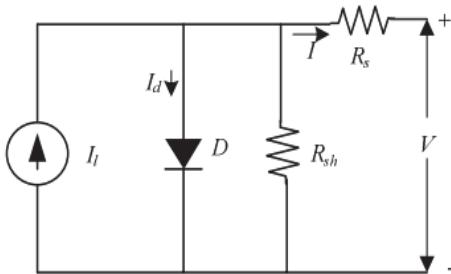


Fig. 2 Corresponding circuit of PV cell

To satisfy the demand, they organized in series and parallel. The sequence assembly increases the voltage of the component, but the parallel assembling raises the array's current. The actual reality of the cell is provided by series resistance (R_s) and parallel resistance (R_{sh}). Utilizing dynamic programming, the solar array is constructed using a mutual series/parallel amalgamation of solar cells to maintain the V-I features. Equation gives the output current of a photo voltaic array. (2)

Where, $I = I_{sc} - I_d$ (2)

$$I_d = I_0 \left(e^{\frac{qV_d}{kT}} - 1 \right) \quad (3)$$

Here, I_0 =reverse saturation current of diode

Q =electron charge by 1.602×10^{-19} coulomb

V_d =voltage across diode

K =Boltzmann constant by $1.38 \times 10^{-23} J/K$

T =junction temperature in Kelvin

From above equations we get

$$I = I_{sc} - I_0 \left(e^{\frac{q(V+IR_s)}{nKT}} - 1 \right) \quad (4)$$

Using suitable approximations,

$$I = I_{sc} - I_0 \left(e^{q((V+IR_s)/nKT)} - 1 \right) \quad (5)$$

The photovoltaic current is I , the PV cell voltage is V , and the diode ideal factor is n . It is critical to understand how insolation and temperature affect PV properties. If the temperature is amplified while the insolation level remains constant, there is a minor growth in cell current but a significant decrease in cell voltage, that is represented in the P-V and V-I characteristics. The PV cell classic is generated in this paper to use a MATLAB script file, and the relevant restrictions are listed in Annex-I.

STATCOM Modeling

Because the electrical power system mostly works by AC measures, and practically all of the loads that are employed require sensitive power, sensitive power reimbursement is a major power eminence issue [7]. The sensitive power flow must be managed to provide the appropriate voltage sustenance for the voltage variation in WECS [5]. Throughout a voltage breakdown, the STATCOM is a benefit in terms of providing more capacitive sensitive power. STATCOMs are power electrical equipment's that can create or absorb sensitive power at their output terminals. If associated to a battery storage device, it is also able to manage real power [12]. To provide sensitive power sustenance to transmission

lines, dissimilar SVCs, it not requires a large number of inductive and capacitive modules [8]. The STATCOM's key advantages are its compact size, which requires less installation space, and its better sensitive power yield at low voltages. From the standpoint of dynamic stability, STATCOM also imparts stronger damping properties [6]. D-STATCOM was used in this paper to improve the power eminence of the hybrid micro-grid. A D-STATCOM attached to the point of common coupling (PCC) could be used to reduce voltage and current-related power eminence problems. When in current control mode, D-STATCOM inserts the harmonic and sensitive modules of the load current to produce the source currents stability and pure sinusoidal. When operating in voltage control mode, the PCC voltage was adjusted having a reference value to protect essential loads from large voltage fluctuations [9].

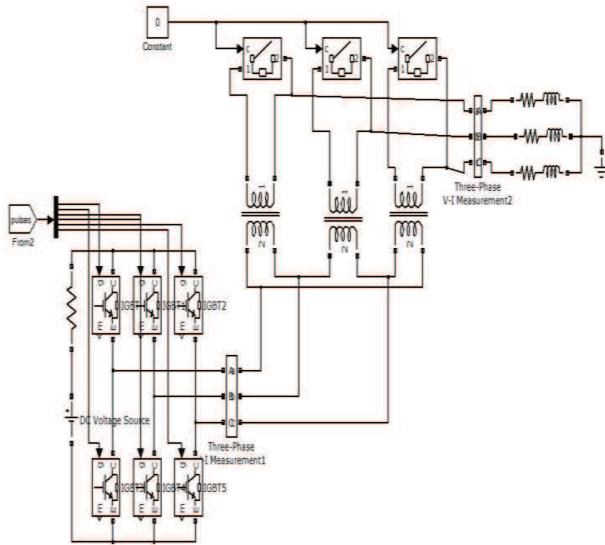


Fig-3 prototype of fuzzy based D-STATCOM

Scheme of Fuzzy Logic Controller

The control system branch suggests a huge controllers used in a variety of industrial and real-time applications. The performance of the system improves once controllers are installed. However, the fundamental issue with traditional controllers is that their performance is greatly reliant on the system's ambient parameters. Due to the complexity of many industrial control systems and their reliance on system parameter fluctuations, a fuzzy method has been shown to be a better solution.

However, rather than totally exchanging the control pattern, the fuzzy pattern is utilized to modify the advances of the SMC controller in dynamic load scenarios. [5].

1. I/O

Here dc link capacitor voltage is synchronized by evaluating the variants in

(i) Variance in dc capacitor voltage

(ii) Altered in change of dc capacitor voltage

$$\text{Diff (K)} = \text{Vdc ref} - \text{Vdc (K)} \dots\dots\dots(6)$$

$$\text{Diff (k)} = \text{vdc ref} - \text{vdc (k)} \quad (6)$$

$$\Delta \text{diff (k)} = \text{diff (k)} - \text{diff (k-1)} \quad (7)$$

Here given equations (6)-(7),

diff (k) and $\Delta \text{diff (k)}$ = is an error with alteration error in k^{th} iteration

$vdc \text{ ref}$ = dc capacitor reference voltage

$vdc(k)$ = dc capacitor voltage of k^{th} iteration

Now the input constraints, fuzzy and SMC controller gain is given by

$$K_p = K_p \text{ ref} + \Delta K_p \quad (8)$$

$$K_i = K_i \text{ ref} + \Delta K_i \quad (9)$$

2. Fuzzification

Fuzzification is the method of transforming real-time corporal values keen on a base anywhere fuzzy action could be performed. It takes the input parameters and assigns a membership value to the membership function that controls the input parameters. Membership functions come in a variety of shapes and sizes, including triangular, sigmoid, and Gaussian for both error and change in error, here are seven dissimilar forms of triangular membership utilities. Figures 4 and 5 depict the input membership functions. The input membership features have been fine-tuned to meet our requirements. Whichever type of membership function it is, it has a membership value allocated to it.

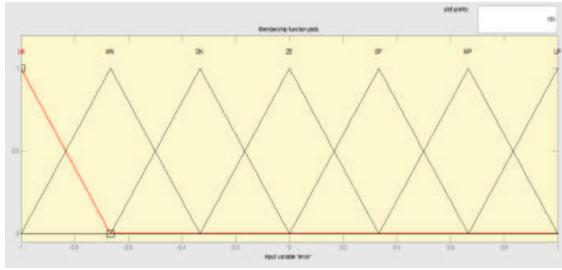


Fig.4. Membership function for error signal

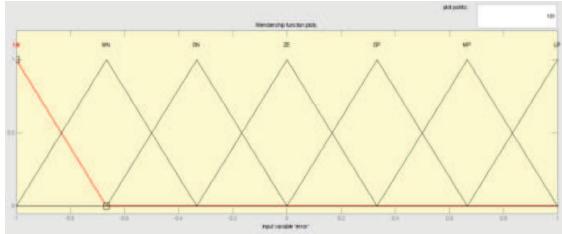


Fig.5. Membership function for change in error signal

3. Implication Appliance

It primarily serves for 2 purposes: a) It determines the law to be applied depended on the membership value allocated to a specific membership function for the present situation; and b) It determines the rule that must be useful depending on the membership value dispensed to a specific membership purpose for the present situation. b) If the rules are stable, the membership value determines the occurrence of a control action. Now, at the conclusion of these steps, we will have a set of rules that may or may not be true. The device action is consequent since a table containing numerous rules, that is commonly referred to as the rule base.

4. Defuzzification

It is the method of changing a fuzzy logic domain to a real-world domain. Basically, it's a process of reverse fuzzification. The definite real world physical values are formed depending on the rules and the principle enumerated fuzzy value. In a nutshell, a fuzzy logic structure is transformed hooked on a human logic structure. Figures 6 and 7 depict the output membership purposes.

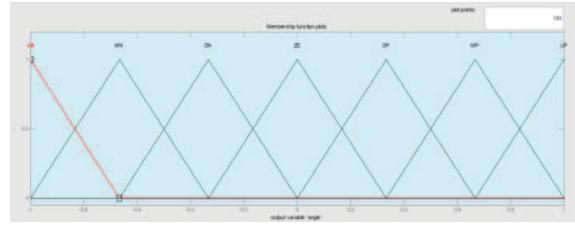


Fig.6. Output Membership function for proportional gain

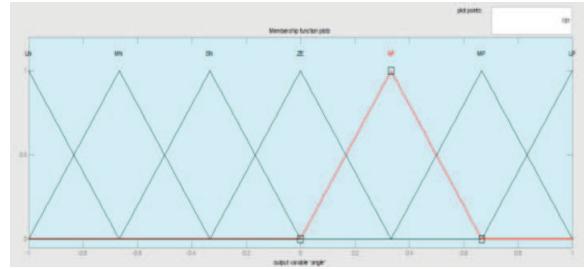


Fig.7. Output Membership function for integral gain

Result And Analysis

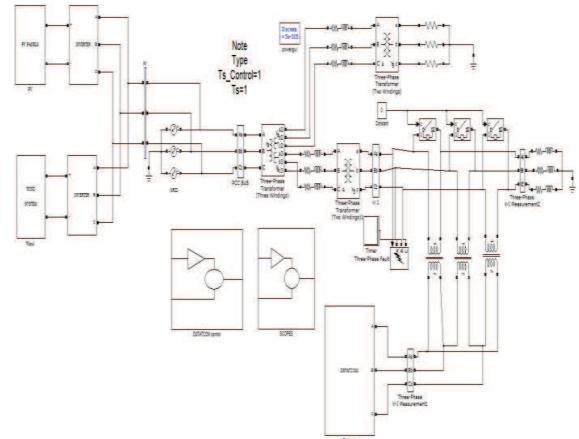


Fig.8. simulation design of DSTATCOM with Fuzzy based PV & wind system

Voltage sag, current harmonics, and power factor are all indicators of power quality. To quantify the power quality delivered by the Wind-PV system, we compute total harmonics distortion (THD). The following equation can be used to compute THD in the current waveform: (3)

Where I_h is the hth-order harmonic current and I_1 is the essential module of current. The THD is determined here using MATLAB's FFT (Fast Fourier Transform) Exploration Toolbox. Computing the THD in the current sent to the grid was used to assess the eminence of the electricity provided by the projected procedure. THD was computed without and with D-STATCOM at a wind speed of 5m/s for a comparative analysis.

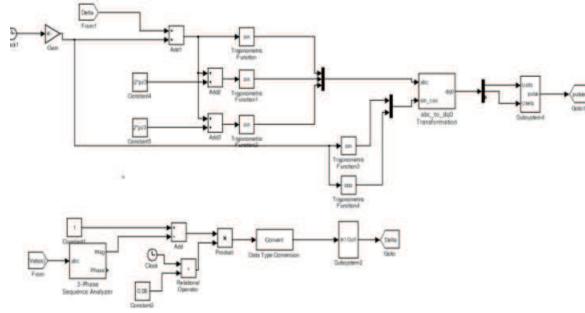


Fig 9: Fuzzy Control circuit

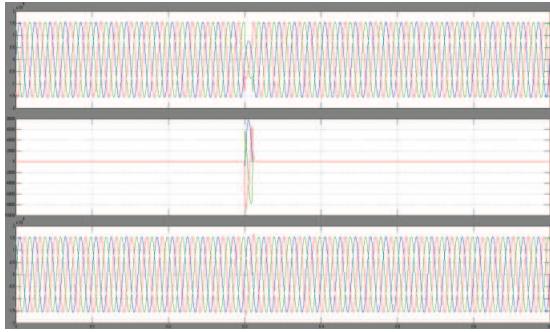


Fig: 10 Sag Voltage, compensation DSTATCOM
Voltage, load voltage

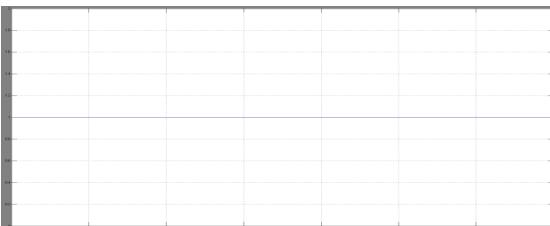


Fig: 11 load power factor (unity)

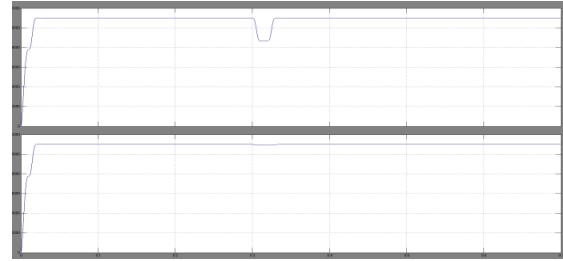


Fig: 12 RMS voltage Sag & load Vrms



Fig: 13 Compensated Active & sensitive powers

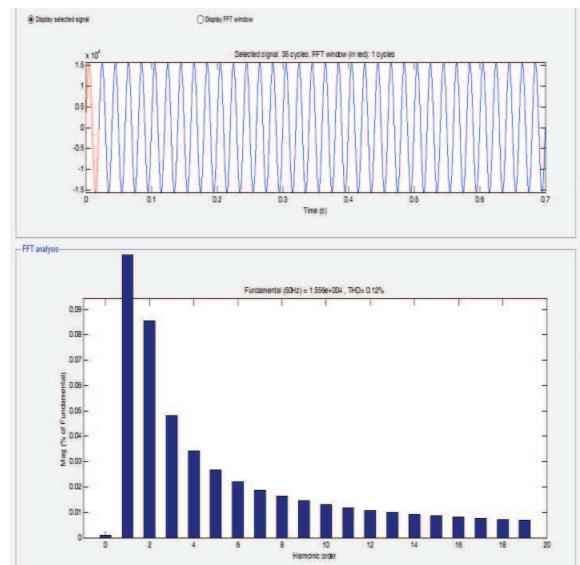


Fig-14 THD in hybrid Wind-PV system with fuzzy based D-STATCOM (T.H.D=0.12%)

S.n o	Compensation	Vrms Drop	Volt age erro r	THD	Power factor
1	Without DSTATCOM	Actual 11kv dropped 5Kv	2kv	15.24	0.6
2	SMC DSTATCOM	Actual 11kv dropped 2Kv	0.8 Kv	1.64	0.8
3	fuzzy DSTATCOM	Actual 11kv dropped 0.1Kv	0.1 Kv	0.12	1 (unity)

Conclusion:

The suggested grid-tied smart PV-WIND-DSTATCOM structure is executed in the laboratory for a 3-phase system to use an adaptive RZA control technique by P & O depending on MPPT method to progress power quality and sustenance the 3-phase AC grid by contributing power to together the grid and the associated loads. The control approach for a 3-phase grid tied single stage PV and wind method by DSTATCOM abilities has been described, which uses an adaptive RZA method to extract reference grid currents and a P&O-MPPT control to maintain max power at all times. The RZA method produced good stable state and transient reactions, and simulation results confirmed the system's dual capability to effort in PV-WIND-DSTATCOM and DSTATCOM methods. The above projected structure works fine above unbalance load and also maintains high power quality.

- [1] G. M. Shafiullah, M. T. Amanullah, A. B. M. Shawkat Ali, P. Wolfs and M. T. Arif, Smart Grids: opportunities, developments and trends, Springer, London, 2013.
- [2] S. J. Steffel, P. R. Caroselli, A. M. Dinkel, J. Q. Liu, R. N. Sackey and N. R. Vadhar, "Integrating solar generation on the electric distribution grid," IEEE Trans. Smart Grid, vol. 3, no. 2, pp. 878-886, June 2012.

[3] T. Wang, D. O'Neill and H. Kamath, "Dynamic control and optimization of distributed energy resources in a micro grid," IEEE Trans. Smart Grid, vol. 6, no. 6, pp. 2884-2894, Nov. 2015.

[4] K. Jager, O. Isabella, A. H. M. Smets, R. A. C. M. M. Van Swaaij and M. Zeman, Solar energy fundamentals, technology and systems, Delft University of Technology, 2014.

[5] T. F. Wu, C. H. Chang, L. C. Lin and C. L. Kuo, "Power loss comparison of single- and two-stage grid-associated photovoltaic systems," IEEE Trans. Energy Convers., vol. 26, no. 2, pp. 707-715, June 2011.

[6] R. Agarwal, I. Hussain and B. Singh, "LMF based control algorithm for single-stage 3-phase grid integrated solar PV system," IEEE Trans. Sus. Energy, vol. 7, no. 4, pp. 1379-1387, Oct. 2016.

[7] Bidyadhar Subudhi and Raseswari Pradhan, "A comparative study on max power point tracking techniques for photovoltaic power systems," IEEE Transactions Sus. Energy, vol.4, no.1, pp.89-98, January 2013

[8] Kandil M, El-Saadawi M, Hassan A, Abo-Al-Ez K. A proposed sensitive power controller for DG grid associated systems." In Proceedings of IEEE international energy conference and exhibition (EnergyCon), IEEE; 2010,pp. 446-51

[9] N.G. Hingorani, Understanding FACTS", IEEE press 1999.

[10] Bogaraj T, Kanakaraj J, "Development of MATLAB/SIMULINK Models for PV and Wind Systems and Review on Control Strategies for Hybrid Energy Systems", International Review on Modelling and Simulations (I.RE.M.O.S.), Vol. 5, 2012.

DenseNet201 for Animal detection and repellent system

Naveenkumar M, AP/CSE

Department of Computer Science and Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
naveenkumar2may94@gmail.com

Nandhakumar B

Department of Computer Science and Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
Nandhakumar.baskar2000@gmail.com

Manoj R

Department of Computer Science and Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
manorc36@gmail.com

Rahul R

Department of Computer Science and Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
rahulsusith@gmail.com

Abstract— Animals entering the agricultural land near the forest areas destroy crops or even attack people residing nearby villages. For humans, agriculture crops, and animals to survive it is essential to utilize technology in a useful way. Therefore, there is a need for a system which detects the presence of animals and produces sound to repel the animals. Our study develops a model utilizing deep convolutional neural networks that detects animals through video analytics. The different features like color, Gabor and LBP are extracted using the segmented images of the animal. Possibilities of combining these features for improving the performance of the detection and classification have also been explored. Detection of animals is accomplished using CNN and symbolic classifiers. For validating the performance of the proposed algorithmic models and also due to non-availability of a large benchmarking related dataset, successful attempts to create an animal image dataset and an animal video dataset. Experimental results show that better detection accuracy is obtained for even partial images of animals. The system will give better accuracy in terms of f1 score precision, recall and it will support the society to safeguard agriculture, farmland, wildlife and human beings.

Keywords—Image processing, Convolutional Neural Network, Animal Detection

I. INTRODUCTION

During crop production in India, nearly 100-200 people and 40-50 animals are killed every year. In spite of deforestation and human intervention in forest areas, the forest environment is not conducive for native animals. Due to the pollution and climatic changes even the forest areas are not receiving normal rainfall in a year. This leads to unavailability of necessary food and water to the animals, so that many animals are coming out of the forest. When wild animals such as tigers, elephants, lions, leopards, bore etc., cause damage to human life, agriculture crops, domestic animals' life, houses, habitat are inevitable [1]. The effect of changes in civilization and lifestyle of human beings are one of the reasons for animals to come out of forest areas. To save animals as well as crops, we need to

repel the animals from the field by monitoring and detecting them.

Nowadays Machine learning has become one of the key fields in the sector of Artificial Intelligence (AI) [2]. Machine learning has features that are able to learn and enhance the performance of the system using previous experience. It doesnot require any specific programming to do so. Programming in machine learning concentrates on accessing data and learning by own using the data. Specific and required patterns in data are identified to make better decisions for improvement of the system in future. Automatic learning, avoidance of human intervention, performance improvement and appropriate adjustments in actions of a system are the key features and characteristics of machine learning.

Deep learning and machine learning act as the subset of Artificial Intelligence. By using the deep learning techniques many performance measures in practical applications such as image recognition, sound recognition etc. are greatly improved when deep learning is applied. Convolutional Neural Networks (CNNs) are algorithms that are widely used to analyze images; they are referred to as deep neural networks in deep learning [3]. Need for pre-processing in CNN is comparatively less over the other image processing algorithms. Hence the concept of CNN along with deep learning is more preferred for image classification applications. It finds a variety of applications such as video and medical image analysis, image recognition, image classifications, recommender systems, and language processing system.

Concept of animal detection is one of the emerging areas in wildlife environment issues. The problem can be solved by watching the activities and pathways of wild animals in forest and its surrounding area by using CNN based animal detection algorithms. This will be helpful to take appropriate actions by the forest department and to safeguard the human being, agriculture crops against attack of animals.

II. LITERATURE

A paper published in 2021, deployed with a cumulative of 1200 images as a dataset from the real world. They used the AI based surveillance system to detect the animals and monitor their presence. They used OpenCv and Pre-trained model Mobile Net SSD for detecting the animals in the farmland. Then the model is trained on MS COCO image dataset. And even though the image is crowded an accuracy of 85% is achieved [4].

In August 2020, this study was conducted in August 2020 utilizing a dataset that consisted of 25 classes of animals, they examined 31,774 images for each class. Several measures are identified to reduce the number of accidents that occur when vehicles collide with wildlife. This will improve road safety and preserve wildlife. With SSD and faster R-CNN, the proposed method has a mAP average and detection speed of 80.07% at 100 frames per second, and 82.25% at 10 fps [5].

A convolution neural network is used in this paper, which was published in 2020, to detect animals even in fragmentary photos. They build their own dataset, which includes animals as well as birds. Animal detection accuracy is 81 percent, while partial image detection accuracy is 47-49 percent. Partial detection is also possible with this approach [6].

In 2021, in this study they prepared a new system is presented that automatically detects and Animals are recognized based on deep learning and segmentation with genetics. Convolutional neural networks help group images given as input by the contributors. The segmentation of images using evolutionary algorithms is proven, as well as this method is unique in terms of precision and accuracy, recall, and MAE. As a result, the proposed approach increases overall results by 99.02 percent precision, 98.79 percent recall, 98.9% F-Measurement, and 0.78% MAE [7].

In this paper, a machine learning technique is used to create an IoT application for animal intrusion. They receive a number of trained datasets for animal recognition and classifications. The R-Cnn algorithm makes precision search to build regions and extracts roughly 2000 areas from each image. The R-CNN gave a mean and accuracy of mAP gives 83.32%, though SSD shows a mAP of 89.42% for creature pictures for 6 classes based on 312 data sets and displays the algorithm's prediction. When compared with other methods, SSD outperforms others in detecting and recognizing the animals [8].

In 2018, Vision - based techniques are employed in this research to develop a low-cost and simple technique for detecting autonomous creatures on roadways, with the purpose of decreasing wildlife collisions. Even though the wild animals are seen at greater speeds, the motorist have not enough chance to avoid an accident. This method achieves a total accuracy of more than 82.5 percent [9].

This paper proposes, with five component identification pipeline is to be used in an IoT based animal recognition system is proposed in this paper. Our technique seems to have a localized mAP of 81.17 percent, a variety and perspective annotated identification efficiency of 94.38 percent and 87.21 percent, etc., and then the AoI gave the efficiency around 73.65 percent for 7 animal targets of interest. Their dataset from WILD includes 5,786 pictures and 12,078 labeled descriptions from 29 categorization types, as well as a number of hard actual detection settings are included in their dataset [10].

Table 1. Summary of various Algorithms and Dataset collection

Source	Data set	Algorithm	Objective
Animal detection in forms using open source computer vision [4]	Own dataset with 1200 images	Mobile net SSD, MS coco image dataset	AI based surveillance system
Animal Recognition and Collision Prevention System using deep learning [5]	Dataset with 31774 images of 25 classes	Based on both SSD and fast R-CNN object Recognition	To avoid collision between humans and animals in high ways
Convolution neural network based animal detection with segmentation [6]	Own dataset with 10 classes	CNN algorithm on the Raspberry Pi3 module	Detection accuracy is obtained even for partial image
A computer system for automatic species detection and recognition [7]	Dataset with 100 distinct subjects of 2 classes	Deep CNN with genetic segmentation	This method provides genetic algorithms with high precision
A machine learning technique for animal intrusion [8]	312 datasets of animal photos for 6 classes	R-CNN and SSD algorithm	R-CNN gave average mean precision and SSD outperforms
Computer vision techniques preventing animal-vehicle collisions [9]	Trained over 2200 images on both positive and negative	computer vision techniques used for autonomous animal detection	Alert has to be triggered once the vehicle speed is upto 36 km/h

III. PROPOSED WORK

A deep convolution neural network based model for detection and classification algorithm is used to detect animals both in videos and images. The detection is performed by getting images/frames from a video camera and pre-trained convolutional neural network is already implemented in raspberry pi with the trained data set. Later the extracted features are fed into a multi-class CNN classifier for the purpose of classification. CNN is constructed using a sequence of layers like Convolutional, subsampling and fully connected Layer. After detecting the wild animal from the speakers connected it makes a crackers sound so that the animals get scared and run away from the farm land. The block diagram shown below describes the overall working process of animal detection system using neural network-based image processing in hardware systems.

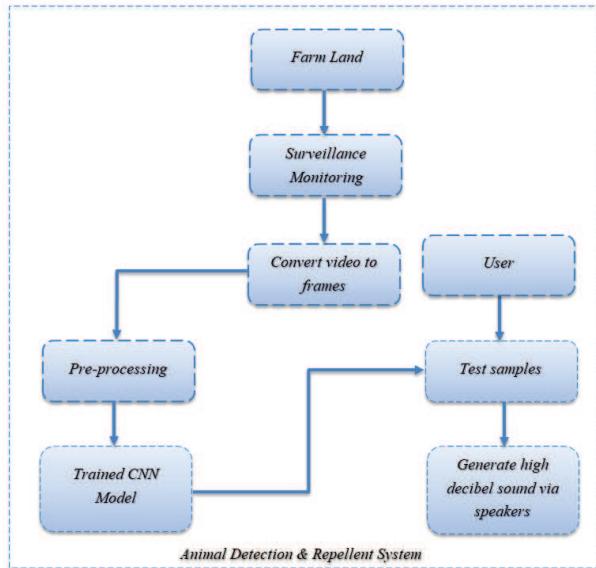


Fig 1 Animal Detection & Repellent System

High resolution video camera is used for capturing the image of animals in the background of nature. Both image and video pictures are able to capture and send the signals to the hardware system. The captured signals from the camera are processed in a given sequence to detect the type of animals. Output signal from the camera is not suitable to process and hence the signals are allowed to pass through the pre-processing system [11]. Afterward, the image frame is processed in a deep convolution neural network model, which is residing in the system.

To get the frames from the streaming or live video we need to install OpenCv and imutils. Then we need to give the path to the pre-trained model so that we initialize a class list and set random color and finally we start the frames per second counter with the help of Video Stream and FPS classes. After getting these frames and sending it to the CNN model with the trained data set. The dataset is collected from Kaggle website

for our project and the same trained through CNN model. Dataset contains 10 different animals with multiple images of animals that differ from each other.

An algorithm is used to build a complete predictive model from fully trained data sets with features and class labels. These predictive models help to estimate the class labels from the new data by using the features from the trained data. The output classes are discrete. In addition to decision trees, Support Vector Machines and many more types of classification algorithms are available.

The CNN is the same as the architecture of the original Le Net (Convolutional Neural Network in Python) and classifies the given picture into various categories. Perceptron machine learning logic is used in CNN with supervised learning mode for analysis of data. Further CNN is mainly constructed by using three layers namely, convolutional, sub-sampling and fully connected layer.

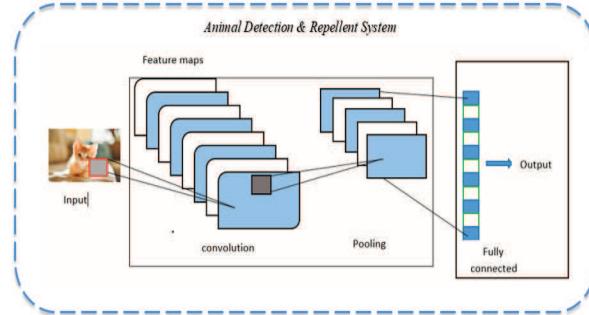


Fig 2 Architecture diagram of CNN model

A. Convolution layer

The Convolution layer has been used at first of CNN layer made up of learnable filters placed spatially a measure of the input layer's height and width. Convolution is processed by sliding every filter across height and width of the input volume and calculating dot product of filter coefficient and input at any position. While doing so for the entire input volume, a two-dimensional activation map is generated which is a response from the filter at each position. Intuitively, a separate 2D activation map will be generated for each filter. At the end, all the entire activation maps are put together and considered as the output of the convolution layer.

B. Pooling Layer

Pooling layers are helpful to decrease the quantity of parameters when the size of the image is huge. Max pooling is applied where the greatest element is considered from the feature map. Dimension of the image is reduced by down sampling technique. Usually filter size of 2x2 is applied for pooling over both height and width and dimension is reduced by 75% while 25% is retained.

C. Fully Connected Layer

Fully Connected Layers are specific types of hidden layer which must be used within the CNN. This is used to combine the features into more attributes that predict the output more accurately. In a fully connected layer, all the neurons are

connected with the previous layer. This is similar to conventional neural networks. Matrix multiplication is applied to determine the activation of neurons.

The proposed animal detection algorithm is developed using convolutional neural network and implemented using a hardware embedded system. Python coding is used to develop the algorithm and implement on Google Co-lab. For the detection of animals with their datasets, the algorithm is run on 2 stages: with training and detecting stages. Calculation is executed in two stages; one is preparing and another is trying for the recognition of creatures. Preparing is done utilizing an enormous informational index containing 1189 pictures arranged through 10 cases for various creatures. In table 1, you can see the informational collection and its cases. For each case, the preparation segment demonstrates the no. of pictures utilized for preparing, while the approval segment demonstrates the quantity of pictures effectively distinguished.

Table 2. Dataset images in validation for various species

Image Data Set	Training	Validation
Tiger	86	75
Lion	80	74
Elephant	83	76
Deer	89	74
Boar	76	70

Since the resolutions of dataset images are high and required resolution for CNN is low, the images are resized to minimize the dimension. Resized pictures are using the training phase and testing phase of algorithm.

IV. RESULT AND DISCUSSION

In recent advancements TPU's and GPU's, there has been a boom in computer vision techniques in the Artificial Intelligence (AI) domain. Google Co-Lab trained the models with GPUs and TPUs over 130 epochs, for each epoch it having n number of steps. Using the preparation dataset the size distributed by the batch size, and the steps are calculated as [12, 13]. In animal detection system we have used the batch size is 32 because it's a basic and popular batch size to train the model faster. The entire dataset is divided with a collection of 32 images and grouped as a single batch likewise the entire dataset is divided equally and trained through the detection system. A detailed description of the various models is provided in the table 4. The model with its performance measure have been listed in table 3.

The webpage is designed for animal detection is shown in Figure 3 it is created by using html and CSS tags for frontend design. TensorFlow JavaScript was deployed for the backend. The Start button is inserted by using the button tag. It gets access to the device's camera after pressing the start button. After that the presence of an animal is detected from the video.



Fig 4 prediction of animal

If the camera detects the presence of an animal, it uses the background code tensorflow.js to identify and classify the species. The accuracy of the species for classification is then given as 1.00 for the best accuracy rate as compared to other species and 0.00 as the worst case with a class of 10 animals.

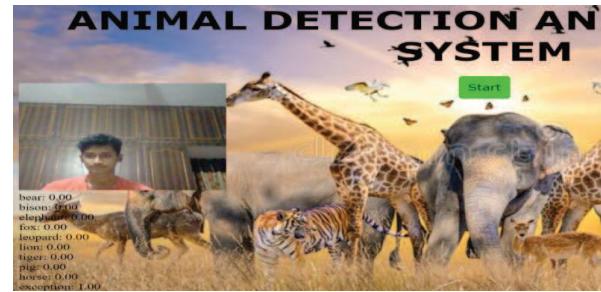


Fig 6 Exception

From the figure 6 it shows that, if the model finds animals and humans that are not in the list of trained datasets, it gets a 1.00. If the model finds species other than those in the list, it is considered an exception. In the updated list, the trained dataset of the model needs to be added to the trained dataset for detecting animals other than those listed on the list. Furthermore, our technique is unique in terms of its grouping of data enhancement and innovative class sample. Image identification would still be a young area, particularly when it comes to camera capture the photos.

Table 3 Comparison with other models

Model	Accuracy	Loss
R-CNN	84.31	15.69
SSD	76.64	23.36
YOLO	71.35	28.65
DenseNet201	93.64	6.36

Table 3 compares different models based on accuracy and loss precision. In this project, we are working with DenseNet201 based on the model size and its parameters. Here the model DenseNet201 is compared with ResNet50 and Inception v3 based on the size of the model,parameters and its

depth. Amount this attributes DenseNet201 have small size of model with more depth in hidden layer to train the model more effectively. When DenseNet is compared to other models such as R-CNN, SSD, and Yolo, it shows that DenseNet has superior accuracy and lower loss precision. The R CNN has an accuracy of 84.31 percent and a loss of 15.69 percent, while the SSD has an accuracy of 76.64 percent and a loss of 23.36 percent. Yolo has a low accuracy and a high loss having (accuracy) 71.35 percent and (loss) 28.65 percent, respectively, when compared to others, DenseNet has a high accuracy of 93.64 percent and a loss of 6.36 percent. In comparison to other models, DenseNet provides superior accuracy and reduced loss.

Our model has led us to a number of results. First, when analyzing trained locations, our model performs well enough for smaller size datasets, with DenseNet201 getting 93.6 percent accuracy.

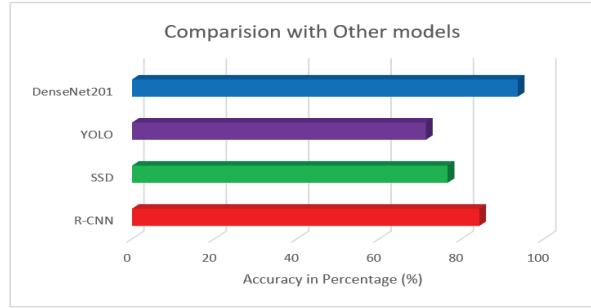


Fig. 7 Performance with other models

DenseNet's performance in comparison to other models is represented through graph in *Figure 7*. The DenseNet201 model outperformed the other algorithms with 93.64 percent because of the model architecture focuses on improving fully connected layers go even deep, it also making it more efficient to train by employing shorter connections between the layers. This indicates that even with a dataset much smaller than those usually reported in this field, significant levels of performance may be achieved using suitable image processing techniques and effective propose solutions. For this study, the ability to apply this technique with minimal data and without shifting cameras across period might be extremely beneficial.



Fig. 8 costing outcome comparison

From the graph in *figure 8*, shows that the DenseNet201 has the least loss percentage because, the Deep CNN is similar to CNN but with more layers. Most basic Approaches have 5 - 10 layers, however the current CNN systems have 20 to 95 layers. When compared to other models, DenseNet201 has the lowest loss percent at 6.36 percent. R-Cnn has the lowest loss percentage at 15.69 percent, SSD has 23.36 percent, and Yolo has the greatest loss percentage at 28.65 percent.

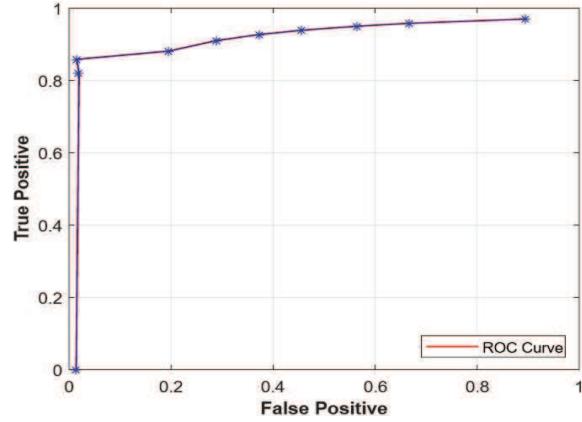


Fig. 9 ROC Curve

Image data were collected in 10 categories with category 10 representing strongest evidence of positivity. For each video frame we can collect 24 cases of images to train the model. Here we can used the 746 sample cases of images with 31 frames and other 2 cases were added. Based on variance-covariance matrix and correlation matrix the summary of the ROC curve is calculated with area and standard deviation of the given area. Here the area is 0.9889 and the standard deviation is 0.0028. The estimated ROC curve with asymmetric confidence interval is shown as graph in figure 9.

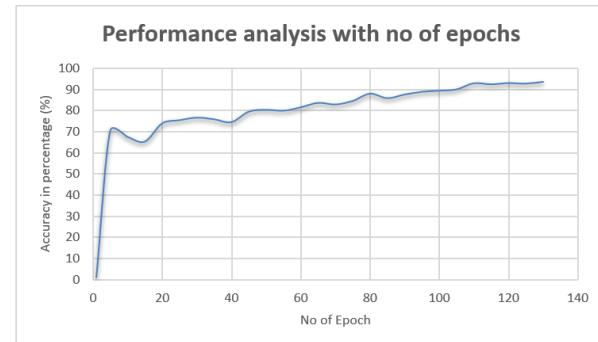


Fig. 10 Performance measure through epochs

When comparing the DenseNet201 model to some other models on epochs 10–130, it is apparent that the DenseNet201 model performs better the other model in terms of loss and accuracy. *Figure 10* graph represents the outcomes of training and validation (evaluation) precision.

Due to night vision images the model take more number of epochs to train the image and give better accuracy only after 100th epochs. For this model we trained up to 140 epochs through GPU. Based on the epochs count we generate the performance measure based on some time interval to predict the accuracy of this animal detection system. At epoch 10 to 30, the model's accuracy graph appears to be kind of overfitting. It can be seen that increasing the epoch value on the DenseNet201 model raises the model's validation accuracy. The future scope of this model is that it may be directly implemented in farmland by making use of hardware modules like Raspberry Pi. Python code can be used for the formalization of the trained datasets. Here the model shows the accuracy of 93.6% with a sensitivity of 93.8% and specificity of 93.3%.

V. CONCLUSION

An efficient animal recognition and repellent system could help the farmers to prevent the wastage of crops and also animal intrusion. A deep learning animal detection model is developed in this study to detect and repel the animals from the farm land. The accuracy for detection and classification of animals is 93.64 percent. The algorithm used for animal detection is convolution neural network (CNN) with the dataset of 10 classes, and the accuracy is measured based on number of epochs for the overall detection. At last, the created model was sent as a web stage to give helpful data to cultivate proprietors by identifying creatures and repulsing them utilizing saltines sound. At this point, this is created as a site page with live location yet. It is normal that these endeavors will add to the government assistance of creatures in the future as an eventual outcome. The proposed strategy can undoubtedly be reached out for discovery of different creatures too after legitimate preparation and testing. The proposed framework can be utilized with other accessible, effective people on foot as a total answer for forestalling creature annihilation.

VI. LIMITATION AND FUTURE SCOPE

In batch size, for any model we can set it as 32, 64, and 128. In this case, our project we used the batch size as 32. But we can also use the batch size of 64 and 128 also. But if we take the batch size as 64 or 128 then the model will take more time for completing one batch at a time. In future large batch size can also be implemented and the same can be compared with this batch size 32. The animal detection system is more wisely used in different farmlands and forecast roads in future.

Research will therefore modify the method of detecting animals to detect animals from all directions and even with partial images in the future. With the right algorithms and designs, the method of detecting and classifying animals will become more accurate.

VII. REFERENCES

- [1] M. Naveenkumar, S. Srithar, V. Vijayaganth and G. Ramesh kalyan, "Identifying the Credentials of Agricultural Seeds in Modern Era", International Journal of Advanced Science and Technology, vol. 29, no. 7s, pp. 4458-4468, 2020.
- [2] V. Vijayaganth; M. Naveenkumar, "4 Smart Sensor Based Prognostication of Cardiac Disease Prediction Using Machine Learning Techniques," in Applications of Machine Learning in Big-Data Analytics and Cloud Computing , River Publishers, 2021, pp.63-80.
- [3] M Naveenkumar et al 2021 J. Phys.: Conf. Ser. 1916 012009
- [4] Rajesh Kumar Bhavani B, K.Mahesh, "Animal detection in farms using open source computer vision", A Journal of Composition Theory, Vol XIV, Issue VII, Jul 2021.
- [5] Saxena A, Gupta DK, Singh S, "An animal detection and collision avoidance system using deep learning", Advances in communication and computational technology. Lecture notes in electrical engineering, vol 668, 2021 . Springer, Singapore.
- [6] Chandrakar, R., Raja, R. & Miri, R. Animal detection based on deep convolutional neural networks with genetic segmentation. *Multimed Tools Appl* (2021).
- [7] K Balakrishna, Fazil Mohammed, C.R. Ullas, C.M. Hema, S.K. Sonakshi, Application of IOT and machine learning in crop protection against animal intrusion, Global Transitions Proceedings, Volume 2, Issue 2, 2021, Pages 169-174, ISSN 2666-285.
- [8] S. U. Sharma and D. J. Shah, "A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique," in IEEE Access, vol. 5, pp. 347-358, 2017.
- [9] Dr. K. Paramasivam, S. Krishnaveni, "CONVOLUTION NEURAL NETWORK BASED ANIMAL DETECTION ALGORITHM FOR PARTIAL IMAGE", AEGAEUM JOURNAL Volume 8, Issue 6, 2020, ISSN NO: 0776-3808.
- [10] J. Parham, C. Stewart, J. Crall, D. Rubenstein, J. Holmberg and T. Berger-Wolf, "An Animal Detection Pipeline for Identification," 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 1075-1083.
- [11] M Naveenkumar and Vishnu Kumar Kaliappan, J. Phys.: Conf. Ser. 1362, pp. 012063, 2019.
- [12] M. Naveenkumar, S. Srithar, B. Rajesh Kumar, S. Alagumuthukrishnan and P. Baskaran, "InceptionResNetV2 for Plant Leaf Disease Classification," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2021, pp. 1161-1167.
- [13] Kavin Kumar D., Vijayaganth V., Naveenkumar M. et.al, "A Cascaded Technique for Hyper Parameters Optimization in Convolutional Neural Network Using K-Means and Bayesian Approach",Journal of Computational and Theoretical Nanoscience, Volume 18, Number 3, March 2021, pp. 833-840(8).
- [14] M. Naveenkumar and Dr. N. Yuvaraj. "Improving the Quality of Service in a Wireless Sensor Network Technologies- A Survey." International Journal for Scientific Research and Development 6.1 (2018): 1851-1856.

An Optimized Gaussian Extreme Learning Machine (GELM) for Predicting the Crop Yield using Soil Factors

R. Prabavathi¹, Balika J chelliah²

¹Research Scholar, ²Associate Professor

^{1,2}SRM Institute of Science and Technology, Ramapuram Campus, Chennai India.

Email: prabavathi.raman@gmail.com

Abstract- Indian agriculture is extremely important and plays a predominant role in economy and employment. The agriculture has seen a significant technological transition because of data collection, environmental factors, crop selection, soil nutrients, pesticides and plant disease for making better farming decisions. This revolution in agriculture is addressed by using emerging technologies. Early detection and management of crop yield indicator problems can help to increase the yield and subsequent profit. Machine learning is an emerging technology used in agricultural research for yield prediction. To produce accurate results, a simplest and very fast optimized learning algorithm called GELM (Gaussian Extreme Learning Machine) classifier with different kinds of activation functions are used. For the soil dataset, the classifier is trained using 50 hidden neurons with different activation functions. The performance analysis of the system shows that gaussian extreme learning achieves an accuracy of 97% compared to other algorithms. This analysis helps in interpretation of results in efficient manner for any regional soil data.

Keywords- Machine Learning, optimized GELM, Soil Nutrients, Precision, Recall, F-Score.

I INTRODUCTION

India is a highly populated country and there is a constant rise of demand in agriculture products. Crops are frequently harmed by extreme weather conditions as a result of global warming. Currently, the agricultural labour force is primarily made up of elderly people who rely on experience and common sense. Furthermore, the lack of adequate equipment to monitor environmental factors, crop selection for the land, soil nutrients, pesticides, and disease types is still a challenge for farmers. This leads to less productivity and quality of the agricultural products. Precision agriculture is the application of an information technology management system to farming. With this system, farmers will have access to valuable information that will help them make better use of their land, increase crop quality and yield, and minimise environmental impacts in order to achieve long-term sustainable farming. This revolution in agriculture is addressed by the recommendation system by an ensemble model. The data is collected based on what type of soil, nutrients required,

duration of crop, seed selection, climatic changes, capacity of water, types of disease, soil parameters etc. An analysis is done on the feature selection and decision making is done for cultivation. The research recommends the famers to select the right crop based on the field parameters observed, thus increasing the yield and also increasing the overall productivity of the nation. The quality and quantity of the food products is also achieved by the system to provide a healthy life to human. The system also emphasizes that the recommendation system provides an efficient and accurate crop production.

II LITERARY SURVEY

In [1] Rahman et al. have used the machine learning algorithms for suggesting the suitable crop for the farmers based on the soil series and type of the land. The experimental results on various algorithms shows that support vector machine (SVM) performs well in suggesting the best crop.

Pawar et al. [2] proposed a system to assist farmers in maximising crop yield. The recommendation system proposed by the authors are based on the fertility and toxicity of the soil using the sensor data. Decision tree J48 algorithm is used which provides good results compared to other classification algorithm.

Kumar et al. [3] have discussed about the problems faced by the farmers in identification of pest and plant disease at the early stage. The proposed method uses SVM, Decision Tree, Logistic Regression in which SVM provides the better results compared to other algorithms.

Kiran et al. [4] proposed a decision-making system that takes soil nutrients, climate, pesticides, and fertilizers as input features. The authors have used classification algorithms for predicting the suitable crop which is based on the dataset collected from 2020-2019. The training and testing of the dataset are based on 70:30 ratio. The observation concludes that the authors have tested on 25 states from which 14 states have produced 99% accurate result, 10 states with 75-99%

accuracy and 1 state with poor result. The results show that the effectiveness of different kinds of algorithms in agriculture.

Sujatha et al. [5] have used 11 soil parameters to classify the soil based on the fertility levels like Low, Medium and High for producing the yield of the crop. The authors have used SVM, DT (decision tree) bagging and boosting, Random Forest and naive bayes algorithm in which RF produces better results.

Malik et al. [6] have done a comprehensive analysis on the atmospheric factors of the soil profile data for predicting the fertility levels of the soil using various machine learning algorithms. The crop factors used for this model are temperature, humidity, pH and rainfall for predicting the yield of the crop like tomato, potato and chilli.

Suchithra et al. [7] have developed a model for the Kerala government using ELM algorithm for predicting the fertility of the soil. The system uses soil factors for soil fertility classification to produce an accuracy of 80%. The authors have analysed using different activation functions.

From the survey, the models proposed by the authors identifies the issues faced by the farmers and provides solution for increasing the yield and maintaining the fertility of the soil. An AI based optimized network models provides digitized solution based on the soil nutrients, pH and climatic patterns for performing effective farming on the lands. The proposed intelligent optimized robust and adaptive ML approach will produce qualitative and quantitative yield production in crops.

III.METHODOLOGY

There are three main challenges the researchers are facing while developing technological solutions to the agricultural dataset. First, varying geographical conditions and complex soil datasets make universal design of prediction algorithms difficult. Second, due to the complexity of the dataset, selection and filtering are critical factors that may result in an underfit/overfit prediction pattern. Third, lack of availability of a universal model makes designing an algorithm to be a challenge in producing qualitative and quantitative yield production of crops. As a result of the growing population, the need for maximum yields using optimal solutions will become a necessity in the near future. A digital agricultural intelligent system that renders reliability and affordability will support the farmers in increasing the yield of a crop.

Therefore, using Gaussian Extreme learning machines for early prediction of soil status will provide an optimal solution and creates a better environment for cultivation.

Data Collection and Pre-processing

Village wise soil data are considered in this system with three different categories. The classification of dataset is based on soil, crop and yield. 15 attributes are considered as input

factors for classification based on the fertility indices as low, medium and high.

Table1: Soil Nutrients with fertility levels

Level Attribute	Low	Medium	High
pH	< 6.5 (Acid)	6.5–7.5 (Neutral)	> 7.5 (Alkaline)
EC	< 1.0 (Non-Saline)	1.0 – 3.0 (Slightly-Saline)	> 3.0 (Saline)
OC	< 0.5	0.5 – 0.75	> 0.75
N	< 280	280 – 450	>450
P	< 11	11 – 22	> 22
K	< 118	118 – 280	> 280
S	< 10	10-15	> 15
Zn	< 1.2	1.2 – 1.8	> 1.8
Fe	< 3.7	3.7 – 8.0	> 8.0
Cu	< 1.2	1.2 – 1.8	> 1.8
Mn	< 2.0	2.0 – 4.0	> 4.0
B	< 0.46	0.46- 1.0	> 1.0

The dataset is collected from the government databases [9][10][11]. The table 1 gives the soil attributes with the fertility indices. The levels of the soil data are indicated as low, medium and high. The fertile property of the cultivated land helps in maintaining the quality of the soil

Extreme Learning Machines

The proposed model is used for predicting the yield of a crop using extreme machine learning algorithm. The soil attributes are collected from the government website. The soil parameters are pre-processed because of the complexity and the nonlinearity of the data.

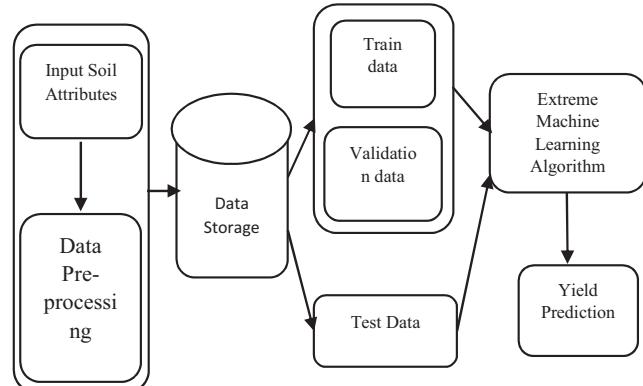


Figure 1: Yield prediction using ELM algorithm

In this model the percentage of training and testing of the dataset is used in the proportion of 80:20. Extreme learning machines will learn more quicker than other learning algorithms. This algorithm is highly efficient at learning neural networks. Therefore, the system uses Gaussian ELM which trains a single hidden layer feedforward network (SLFN) with various activation functions for achieving the lowest training error.

IV. PERFORMANCE AND RESULTS

The Extreme learning machines has outstanding performance results on the real-time applications which are used for classification, clustering and regression. Accuracy, kappa, precision, recall, and F-Score are the five parameters used to calculate prediction performance [8].

$$Accuracy = \frac{\text{Number of Correct predictions}}{\text{Total Predictions}} \quad (1)$$

$$Kappa = \frac{\text{observed accuracy} - \text{expected accuracy}}{1 - \text{expected accuracy}} \quad (2)$$

$$Precision = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})} \quad (3)$$

$$Recall = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})} \quad (4)$$

$$F\text{-score} = \frac{2 * (\text{precision} * \text{Recall})}{(\text{precision} + \text{Recall})} \quad (5)$$

Where TP - True Positives FP- False Positives

FN- False Negatives TN-True Negatives.

Different types of soil attributes are considered to find the yield of the crop and performance is measured using.

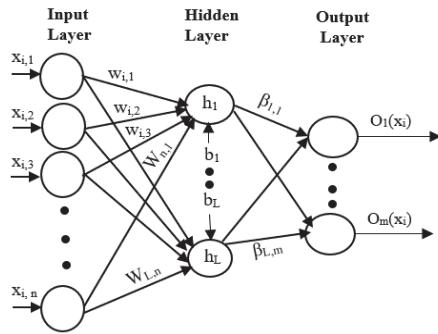


Figure 2: Extreme Learning Machines

Figure 2 shows the architecture of extreme learning machines. It is made up of three layers: the input layer, the hidden layer, and the output layer.

The input vectors denotes x_1, \dots, x_n and output vector denotes o_1, \dots, o_m . The input weights are denoted by w_i and biases is

denoted by b_i . The output function of a single hidden layer feedforward network (SLFN) is denoted as

$$f(x) = \sum_{i=1}^L (\beta_i h_i(x)) \quad (6)$$

The activation functions are not differential in GELM. Approximation of target functions $f(x)$ can be done by tuning the parameters of the hidden nodes. The GELM algorithm is a simple and fast learning algorithm whose performance with different activation functions predicts the fertility levels using the soil factors. In this system hyperbolic tangent function, gaussian radial basis and sine squared function are considered as the activation function for predicting the soil nutrient fertility index.

1	N	P	K	temperature	humidity	ph	rainfall	label
2	90	42	43	20.87974	82.00274	6.502985	202.9355	rice
3	85	58	41	21.77046	80.31964	7.038096	226.6555	rice
4	60	55	44	23.00446	82.32076	7.840207	263.9642	rice
5	74	35	40	26.4911	80.15836	6.980401	242.864	rice
6	78	42	42	20.13017	81.60487	7.628473	262.7173	rice
7	69	37	42	23.05805	83.37012	7.073454	251.055	rice
8	69	55	38	22.70984	82.63941	5.700806	271.3249	rice
9	91	53	40	26.52724	81.41754	5.386168	264.6149	rice
10	99	50	15	18.1471	71.09445	5.573286	88.07754	maize
11	74	56	22	18.28362	66.65953	6.829199	80.97573	maize
12	83	45	21	18.83344	58.75082	5.716223	79.75329	maize
13	100	48	16	25.71896	67.22191	5.549902	74.51491	maize
14	79	51	16	25.33798	68.49836	5.586245	96.4638	maize
15	94	39	18	23.89115	57.48776	5.893093	102.8302	maize
16	20	68	17	30.11873	60.11681	6.578715	71.7298	blackgram
17	43	68	20	29.57813	66.17588	7.497469	69.43895	blackgram
18	44	76	22	27.26459	68.01233	7.775306	68.91754	blackgram
19	34	60	16	31.35731	64.24992	7.322555	63.85669	blackgram
20	21	72	17	31.52105	66.55724	7.580527	61.71111	blackgram

Figure 3: Soil dataset

Figure 3 shows the dataset collected for producing good yield prediction. For implementation python programming language is used because of its efficiency and there are more library functions which supports machine learning. Here NumPy and Pandas are used for data analysis and visualization.

Table 2: Performance Results

Algorithm	Precision	Recall	F-Score	Accuracy
Decision Tree	0.86	0.90	0.87	0.90
Random Forest	0.90	1.00	0.92	0.95
SVM	0.87	0.88	0.85	0.86
Naïve Bayes	0.92	0.94	0.95	0.94
ELM	1.00	0.95	0.98	0.97

The prediction of crops produces an accuracy of 97% using GELM algorithm. The proposed work is compared with other algorithms like SVM, Naïve bayes, Random Forest and decision trees. The confusion matrix is obtained for all the algorithms. Table2 shows the performance results of the algorithms. The comparative analysis of algorithms used for the prediction of the yield of the crop shows that GELM produces good results. The accuracy of all the algorithms for the input dataset is shown in figure 4.

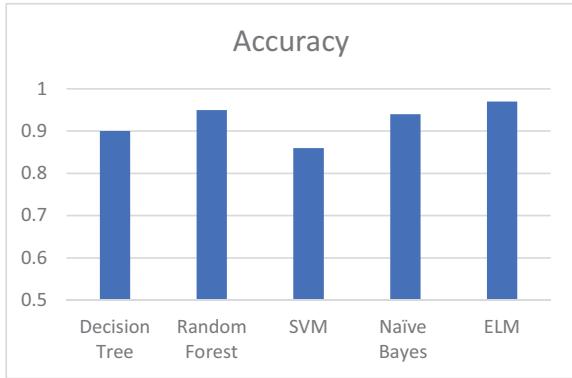


Figure 4: Yield Prediction

The ELM algorithm produces good accuracy compared to all other algorithms. Figure 5 shows the performance metrics of crops.

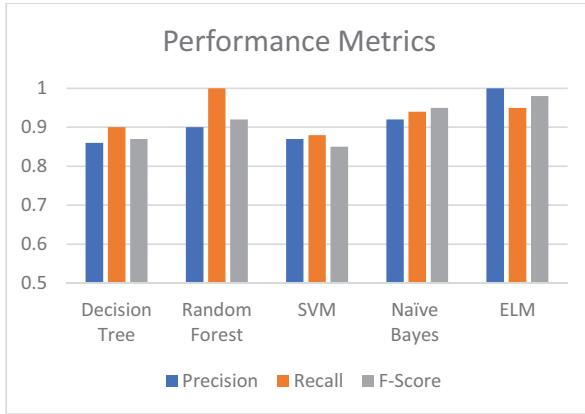


Figure 5: Accuracy of machine learning algorithms

From the performance metrics it shows that the algorithm is efficient on any variety of crops chosen based on the soil parameters.

VII. CONCLUSION

Soil nutrient analysis is one of the biggest challenges faced by the researchers in producing the yield of the crop. In this proposed work, an analysis of soil nutrients is done using different machine learning algorithms like SVM, Naïve Bayes, Neural net, Decision Tree and ELM in which gaussian extreme machine learning algorithm produces the best results compared to other algorithms. Based on the performance metrics GELM produces 97% accuracy and its efficient compared to other algorithms. This digital solution helps the farmers and Tamil Nadu government to produce improved yield production and soil quality. In future this system can be used in any geographical areas with micro nutrients.

REFERENCES

- [1] S.A.Z.Rahman, K.Chandra Mitra and S.M.Mohidul Islam, "Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series", 21st International Conference of Computer and Information Technology (ICCIT), 2018, pp.1-4.
- [2] M.Pawar and G. Chillarge, "Soil Toxicity Prediction and Recommendation System Using Data Mining in Precision Agriculture", Third International Conference for Convergence in Technology, 2018, pp. 1-5.
- [3] A.Kumar, S.Sarkar and C.Pradhan, "Recommendation System for Crop Identification and Pest Control Technique in Agriculture", International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0185-0189.
- [4] K. M P and D. N R, "Crop Prediction Based on Influencing Parameters for Different States in India- The Data Mining Approach", Fifth International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1785-1791.
- [5] S. M and J. C D, "Classification of Soil Fertility using Machine Learning-based Classifier", Second International Conference on Secure Cyber Computing and Communications (ICSCCC), 2021, pp. 138-143.
- [6] P.Malik, S.Sengupta and J.S.Jadon, "Comparative Analysis of Soil Properties to Predict Fertility and Crop Yield using Machine Learning Algorithms", Eleventh International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 1004-1007.
- [7] M.S.Suchithra and Maya L, "Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters", Information processing in Agriculture, Elsevier, Vol 7, 2020, pp.72–82, 2020.
- [8] David MW, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation", arXiv preprint, 2020
- [9] TNAU Agritech portal <http://agritech.tnau.ac.in/>
- [10] Open Government data: <https://data.gov.in/>
- [11] Indian Soil Health Card data. <https://soilhealth.dac.gov.in/>.

Maneuvering Control of Autonomous Underwater Vehicle (AUV) by Simulated Annealing and Moth Flame Optimization Tuned Controllers

Sudarshan K. Valluru

*Center for Control of Dynamical Systems and Computation
Dept. of Electrical Engineering
Delhi Technological University
Delhi-110042, India
sudarshan_valluru@dce.ac.in*

Hitesh Thareja

*Center for Control of Dynamical Systems and Computation
Dept. of Electrical Engineering
Delhi Technological University
Delhi-110042, India
hiteshthareja_2k18ee084@dtu.ac.in*

Abstract—This paper emphasizes maneuvering control of Autonomous Underwater Vehicle (AUV) with the help of Linear-Proportional Integral Derivative (L-PID) and Fractional Order PID (FOPID) controllers. The values of the gain parameters of L-PID (K_p, K_i, K_d) and FOPID ($K_p, K_i, K_d, \lambda, \mu$) are tuned by Simulated Annealing (SA) and Moth Flame optimization (MFO) techniques. The performance of both the controllers is compared and the better algorithm to tune the parameters of the controllers is determined. A step function is applied to the system's input to test the AUV, and the corresponding response to the input function is examined. The transient parameters of the AUV, such as rise time, settling time, overshoot, and performance index ITAE were calculated and compared. The investigatory results show that both the algorithms used to tune the FOPID controller accurately steer the AUV. Also, the MFO tuned FOPID controller is more efficient than the SA tuned FOPID controller.

Keywords— *AUV, LPID, FOPID, Moth-flame Optimization, Simulated Annealing*

I. INTRODUCTION

Autonomous Underwater Vehicle is a system that can accomplish many tasks under the water. The vehicle proves to be an instrumental technology for various departments concerning military operations, commercial and research purposes, etc. In the deep sea, AUV proves to help gather information about the environment where human reach is difficult. It enables the vehicle to be controlled in harsh surroundings without any human interference. The robust control of the AUV is difficult due to the more subsystems embedded into it. The six degrees of freedom (DoF) ensure the consistency of the vehicle. A pictorial representation of the AUV[1] is depicted in Fig.1, and the notations [2] are given in Table. I. The AUV is a quite complex system, due to the natural and unexpected disturbances in the water, thereby the maneuvering control of AUV becomes a complex task, yet.

To resolve this complex task, researchers have tried many intelligent techniques[3], [4] to control the AUV under natural and unexpected disturbances more precisely. But, it seems that controlling an AUV with controllers like PD, PID[5] is a more straightforward approach with satisfactory results. Additionally, this type of scheme appears to be an easy-to-use application for linearized systems. In contrast, however, PID controllers[6] suffer from tedious calculations during changes in system parameters due to the occurrence of natural perturbations. Many research papers have also been proposed in the literature for the application and control of the system, using the LPID and FOPID controllers. Such controllers are also required to control many other real-time systems, such as tracking and control of the helicopter[7], cart pendulum[8], and Ball beam[9] benchmarked systems.

For the robust control of an AUV, this experimental study formulates LPID and FOPID controllers that are effectively tuned using a bio-inspired meta-heuristic optimization algorithm named the Moth-Flame Optimization (MFO) and Simulated Annealing (SA). The performance of the FOPID controller is found to be superior to that of the LPID controller, and the MFO algorithm is found to be more reliable than the SA algorithm in maneuvering control of the AUV.

The arrangement of the topics in this paper is AUV system is analytically modelled in Section II. After that, Section III contains the design of LPID and FOPID controllers. Section IV involves designing and implementing LPID and FOPID controllers using MFO and SA algorithms. After that, section V shows the experimental results and responses obtained. Finally, section VI ends with the conclusion.

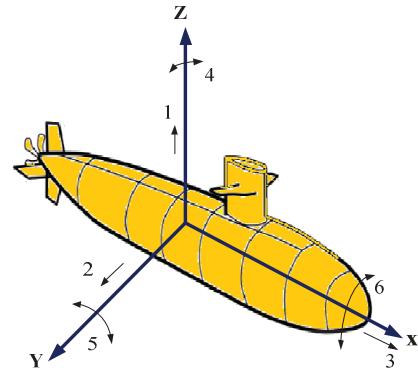


Fig1. Six DOF of an AUV

TABLE I. NOTATIONS OF AUV

Direction Of Motion	Force & Moment	Fixed Earth Frame (Position)	Fixed-Body Frame(velocity)
Surge (Motion along X – axis)	X	x	u
Sway Motion along Y – axis)	Y	y	v
Heave (Motion along Z-axis)	Z	z	w
Roll (Rotation along X – axis)	K	ϕ	p
Pitch (Rotation along Y–axis)	M	θ	q
Yaw (Rotation along Z – axis)	N	ψ	r

II. AUV MODELLING

The first step in analyzing the AUV is calculating the transfer function and stability. The transfer function would then be used to design the controllers for maneuvering control of the AUV. The system's transfer function provides us information about the yaw angle, the deflection parameter, and a relation between the above parameters.

Thus, an AUV's mathematical modelling is proceeded by examining two separate reference frames, known as the earth-fixed and body-fixed frames of reference. AUV coordinates are defined by considering the three axes perpendicular to each other, starting at an arbitrary location. The X-axis shows a movement direction, while Y-axis corresponds to the orthogonal direction and Z-axis depicts the depth. Equations (1) and (2) describe the velocity vector (v) and position vector (η) of an AUV respectively.

$$v = [u \ v \ w \ p \ q \ r]^T \quad (1)$$

$$\eta = [x \ y \ z \ \phi \ \theta \ \psi]^T \quad (2)$$

On simplifying the equations (1) and (2) in the pure steering plane and considering the origin of the body-fixed frame to synchronize with the center of gravity:

$$m(v + mu_0\dot{r}) = \sum Y \quad (3)$$

$$I_Z\ddot{r} = \sum N \quad (4)$$

The surge speed (u_0) value is assumed, and the value has been set at 0.75 m/s. To simplify our calculations, it considered that pitch and roll angles to be minimal and thus can be neglected:

$$\Psi = \frac{\sin\phi}{\cos\theta}q + \frac{\cos\phi}{\cos\theta}r \approx r \quad (5)$$

Equation (6) can be deduced in the matrix form by rewriting the equations from (1) to (5).

$$\begin{bmatrix} m - Y_v & -Y_r & 0 \\ -N_v & I_{ZZ} - N_r & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{r} \\ \dot{\Psi} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} v \\ r \\ \Psi \end{bmatrix} = \begin{bmatrix} Y_{\delta_i} \\ N_{\delta_i} \\ 0 \end{bmatrix} \delta_r \quad (6)$$

Values of the corresponding parameters of the AUV[10] are substituted in the above equation to get the state equations (7) to (9)

$$\dot{x}(t) = Cx(t) + Du(t) \quad (7)$$

Where

$$C = \begin{bmatrix} -0.114 & -0.2647 & 0 \\ 0.0225 & -0.2331 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$D = [0.0211 \ -0.0258 \ 0] T$$

$$U = \delta_r \quad (8)$$

Equation (9) shows the association of the yaw angle (ψ) for the rudder deflection (δ_r).

$$\frac{\Psi(s)}{\delta_r(s)} = \frac{-0.0258s - 0.0024}{s^3 + 0.3445s^2 + 0.0319s} \quad (9)$$

which is used for designing controllers for the AUV.

III. DESIGN OF CONTROLLERS FOR AUV

This section is dedicated to designing the controllers for efficient controlling of the AUV. The two controllers proposed to control the AUV are LPID and FOPID controllers.

A. LPID CONTROLLER

The AUV maneuvering control LPID is a conventional control scheme based on the three control parameters (k_p, k_i, k_d). The equation for LPID controller with unity feedback is depicted as equations (10), (11), and (12).

$$c(t) = k_p e(t) + k_i D^{-1}e(t) + k_d De(t) \quad (10)$$

$$U(s) = \frac{c(s)}{e(s)} = k_p + k_i s^{-1} + k_d s \quad (11)$$

$$1 + G(s)U(s) = 0 \quad (12)$$

After obtaining the feedback equation, the next task is to find the optimal gains of the three parameters of the LPID controller (k_p, k_i, k_d). This is achieved by optimizing the characteristic equation, thus minimizing the error based on ITAE- Integral of the Time multiplied by Absolute Error.

$$ITAE = \int_0^t t|e(t)|dt \quad (13)$$

The schematic diagram of MFO/SA tuned LPID Controlled AUV is shown in Fig.2.

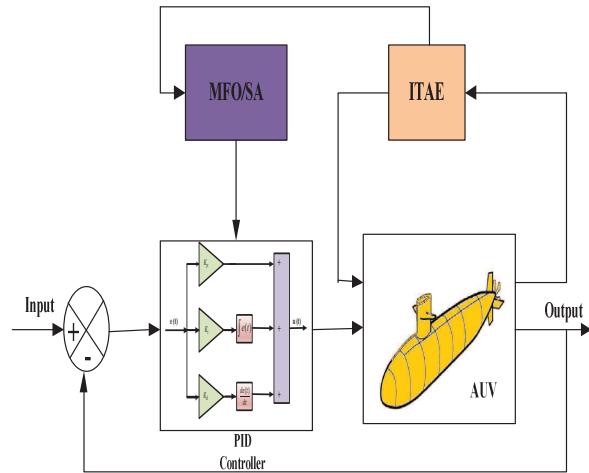


Fig.2. Schematic diagram of the feedback system of LPID on AUV

B. FOPID CONTROLLER

FOPID is another type of controller under the PID family which consists of five controlling parameters ($k_p, k_i, k_d, \lambda, \mu$), thus, the AUV response is more accurate than the LPID controller. The characteristic equation of a FOPID controller[11] and the unity feedback system using the controller are shown in equations (14) and (15).

$$U(s) = \frac{c(s)}{e(s)} = k_p + k_i s^{-\lambda} + k_d s^{\mu} \quad (14)$$

$$1 + G(s)U(s) = 0 \quad (15)$$

The MFO/SA tuned FOPID for AUV is shown in Fig. 3.

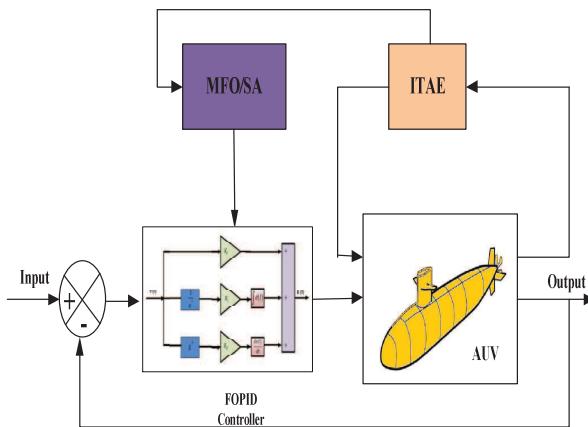


Fig.3. Block Diagram of the feedback system of FOPID on AUV

IV. IMPLEMENTATION OF SA AND MFO FOR TUNING THE CONTROLLERS

Tuning is defined as obtaining the controller's parameters to precisely controlling our system. LPID and FOPID controllers can be very useful in controlling the dynamical systems, but the tuning of these controllers requires several iterations and computational techniques to achieve our desired result. They also need constant adjustment of the control approach to avoid the system's stagnant reaction. Various types of algorithms can be used to optimize the controller efficiently. One class of the algorithm based on bio-inspired techniques is discussed. The bio-inspired algorithms can be used to minimize the struggles in the tedious tuning of the controllers and achieve the desired results.

A. Simulated Annealing Tuned Controllers

One of the most extensively utilized approaches for solving the challenges of dynamical systems is Simulated Annealing. This simulated annealing[12] method is inspired by the response of the actual scenario in the heating environment and is based on statistical mechanics theories. A material kept in a heating environment is allowed to cool very slowly in steps till it reaches a state of lowest energy or optimized state.

The pseudocode of the SA algorithm to tune the gain parameters of LPID/FOPID controllers is :

Step 1: The gain parameters of an LPID/FOPID controller have fixed ranges defined by an objective function $f(x) = [K_p, K_i, K_d]^T$ or $[K_p, K_i, K_d, \lambda, \mu]^T$

Step 2: Consider a temperature (t_0) at the beginning

Step 3: Initially, s_i is taken as the state of the system

Step 4: The lowest energy or optimal state is represented by s_{opt}

Step 5: At the beginning of the process set t as t_0 and s_{opt} as s_i

Step 6: **for** $t = 1$ to t_{max} **do**

```

Assign  $s_{opt+1}$  to adjacent/nearest( $s_{opt}$ )
Change  $\Delta E$  to  $(f(s_{opt+1}) - f(s_{opt}))$ 
if  $\min(1, e^{-\Delta E/t}) \geq \text{random}(0,1)$  then
    Update  $s_{opt}$  to  $s_{opt+1}$ 
end if
    update the value of t with the temperature schedule (t)
end for
Step 7: The final optimized state can be presented as output.

```

The gain coefficients for LPID controller K_p , K_i , and K_d derived with the help of SA techniques are -19.9298, -4.9419, and -41.25, respectively, and the gain coefficients for FOPID controller K_p , K_i , K_d , λ , μ obtained using the SA method are -18.296, -1.9659, -41.7380, 0.9187 and 1.0219 respectively. Simulated Annealing is an efficient algorithm even though it exhibits a significant disadvantage during iterations in course of minimizing the cost function (ITAE). Still, in some cases, the values of parameters also get accepted where the cost function 'ITAE' increases instead of decreasing. This number of cases occurs with a probability denoted by p called the transition probability. To overcome this disadvantage, MFO is implemented to realize the AUV more efficiently.

B. Moths Flame Optimization Algorithm tuned Controllers

MFO is a newly designed mechanism based on the nature of the movement of moths. It uses the principle of population search to find the optimized solution of a cost function. This bio-inspired optimization technique is exceptionally adaptable and may be used to identify the best answer to various real-world issues. Mirjalili [13] introduced the moth-flame optimization (MFO) approach. It starts by randomly producing moths inside the solution zone using a transverse orientation, as illustrated in Fig.4.

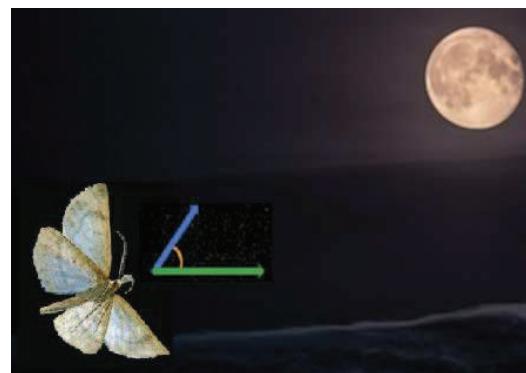


Fig.4 Traversal of Moth

The value of the fitness function for each moth is then computed, and flame tags the ideal spot. The moths' locations are then updated on the spiral movement function to achieve appropriate positions ordered according to a flame. The current acceptable locations of individuals are updated, and the previous actions are replicated. The process continues for a defined value of iteration.

There are three main postulates in the MFO algorithm. These are discussed below:

1. Population Initiation:

It is assumed that the moths have the freedom to fly in all dimensions. It is represented by the matrix form in equation(16).

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & \dots & b_{1,d} \\ b_{2,1} & \dots & \dots & \dots & b_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{x,1} & b_{x,2} & \dots & \dots & b_{x,d} \end{bmatrix} \quad (16)$$

x denotes the number of moths

The number of dimensions is denoted by d . These dimensions remain in the region of solution.

Following is an array that shows the fitness function at each step:

$$OB = \begin{bmatrix} OB_1 \\ OB_2 \\ \vdots \\ OB_x \end{bmatrix} \quad (17)$$

The d -dimensional space contains the flames that are represented by the following matrix:

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & \dots & \dots & F_{1,d} \\ F_{2,1} & \dots & \dots & \dots & F_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ F_{x,1} & F_{x,2} & \dots & \dots & F_{x,d} \end{bmatrix} \quad (18)$$

$$OF = \begin{bmatrix} OF_1 \\ OF_2 \\ \vdots \\ OF_x \end{bmatrix} \quad (19)$$

Moths and flames are the solutions. Each iteration handles and updates the main distinction between moth and flames. Moths are the actual searching agents that travel about the search area. Flames are the best places for moths that have been discovered so far.

2. Updation in the Position of Moths:

The below three steps of the algorithm helps us to find the global value of a problem.

$$MFO = (I, P, T) \quad (20)$$

In the equation(20), I denote the initial moth in the population arbitrarily. $I: \phi \rightarrow \{B, OB\}$. The movement of moths within the constrained area is denoted by, $P: B \rightarrow B$. To end the process, a criterion is to be mentioned, denoted by $T: B \rightarrow \text{true}, \text{false}$

The equation (21) explains function I , which applies random distribution.

$$M(c, d) = (ub(c) - lb(d)) * \text{rand}(\cdot) + lb(c) \quad (21)$$

Where lb , ub are upper bound and lower bound respectively. When using a logarithmic spiral, there are three things to keep in mind are:

- The spiral must commence with the moth as its beginning point.
- The spiral's final point must be at the flame position.
- The range variation of the spiral should not go beyond our search space.

$$S(B_c, F_d) = D_c \cdot e^{bt} \cdot \cos(2\pi t) + F_d \quad (22)$$

D_c denotes the region between c^{th} moth and d^{th} flame.

$$Dc = |Fd - Bc| \quad (23)$$

The shape of the logarithmic spiral is described by Bp . Also, t represents a random value within the range $[r, 1]$. Close to the flame, the spiral motion balances exploitation and exploration. The variable r has a range $[-1, 2]$ during the iterations in the algorithm. It is termed the convergence constant. Fig.5 shows a pictorial view of the facts mentioned above.

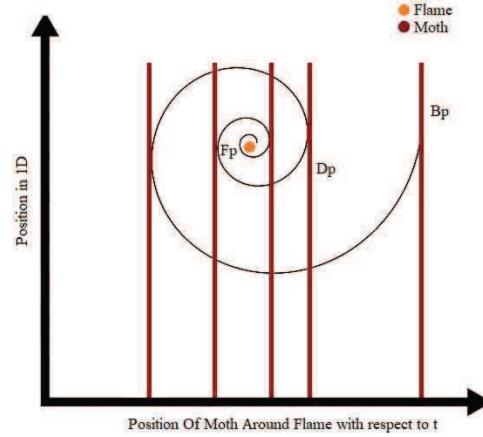


Fig.5. Logarithmic spiral shape

3. Flame Updation

Moth locations are modified at n various locations around the search zone, limiting the best solutions. As a result, using the equation (24), reducing the flames resolves the conflict:

$$\text{flame no.} = \text{round} \left(N - l * \frac{N-1}{T} \right) \quad (24)$$

Total flames are represented by N

Present iteration is represented l

Total iterations are represented by T

The mathematical analysis of the MFO in the form of pseudocode is:

Step 1: Range for control parameters is assigned as

$$f(x) = [K_p, K_i, K_d]^t \text{ or } [K_p, K_i, K_d, \lambda, \mu]^t$$

Step 2: The population of Moth-Flame is defined

Step 3: Moth B 's position is initialized randomly

Step 4: **for** $p = 1$ to n **do**

 Evaluate the value of fitness function F

end for

Step 5: $i =$ iterations

While $I \leqslant$ total i **do**

 Eq. (23) is used to update flames

$OB = \text{Fitness Func}(B)$;

if $i == 1$

$F = \text{sort}(B)$;

$OF = \text{sort}(OB)$;

else

$F = \text{sort}(B_{t-1}, B_t)$;

$OF = \text{sort}(B_{t-1}, B_t)$;

end

Step 6: **for** $p = 1:n$

for $q = 1:d$

 The values of r and t are updated

 Eq. (22) can be used to find the value of D

 Eq. (21) is used to update $S(p, q)$

end

end

Step 7: Finally, obtained the minimized cost function with efficiently tuned parameters

The parameters obtained after tuning the controllers with the MFO algorithm for LPID, the gain parameters K_p , K_i , and K_d are calculated as -15.5343, -0.025, and -38.93, respectively. For the FOPID controller, the gain parameters K_p , K_i , K_d , λ , and μ are calculated as -13.5283, -0.0261, -40.2255, 0.4405, and 0.9987, respectively.

V. SIMULATION AND RESULTS

The unit step is provided to the AUV controlled by the LPID and FOPID controllers. The response of the AUV is recorded to compare the performance of controllers.

The tuned controller's response for SA-LPID and SA-FOPID is shown in Fig.6. Whereas MFO-LPID, and MFO-FOPID controllers for AUV maneuvering control is shown in Fig.7. Also, a comparative depiction between MFO tuned FOPID and SA tuned FOPID is represented in Fig.8.

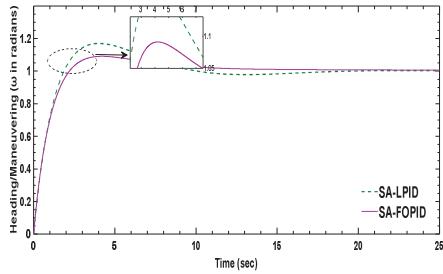


Fig.6 Response for SA tuned FOPID and LPID controllers

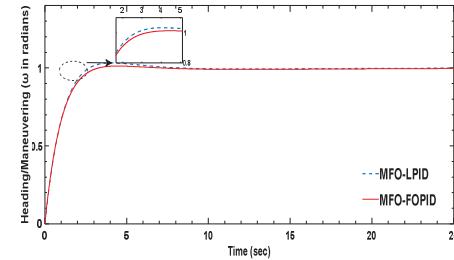


Fig.7. Response for MFO tuned FOPID and LPID controllers

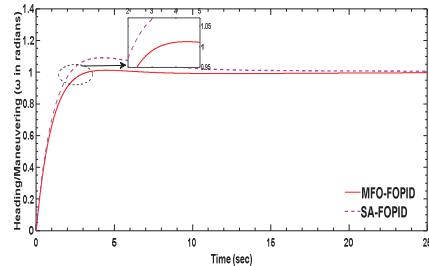


Fig.8. Response for MFO and SA tuned FOPID controllers

Table II represents SA tuned gain values for LPID and FOPID controlled AUV. Table III depicts MFO tuned gain parameters of LPID and FOPID controllers.

TABLE II. SA TUNED GAIN COEFFICIENTS

Parameters	Gain Coefficients using SA algorithm	
	LPID	FOPID
K_p	-19.9298	-18.2960
K_i	-4.9419	-1.9659
K_d	-41.25	-41.7380
λ	-	0.9187
μ	-	1.0219

TABLE III. MFO TUNED GAIN COEFFICIENTS

Parameters	Gain Coefficients using MFO algorithm	
	LPID	FOPID
K_p	-15.5343	-13.5283
K_i	-0.025	-0.0261
K_d	-38.93	-40.2255
λ	-	0.4405
μ	-	0.9987

The transient characteristics of the LPID controller with the utilization of MFO and SA are compared in Table IV.

TABLE IV. TUNED CONTROLLER TRANSIENT PERFORMANCE

Controller	ITAE			
	LPID		FOPID	
Tuning Algorithm	SA	MFO	SA	MFO
Rise Time(Sec)	1.3813	1.7334	1.5418	1.834
%Overshoot	16.9740	3.1643	8.885	1.264
Settling Time(Sec)	14.1150	5.61	10.5802	2.7768
ITAE	3.24	2.72	2.06	1.67

Table IV shows that when the AUV is controlled by the FOPID controller using a particular algorithm has much better performance than the LPID controller. In addition to this, it also observed that MFO tuned FOPID controller shows superior performance than SA tuned FOPID controller.

VI. CONCLUSION

This paper intends to provide an insight into the comparison of the two controllers used to control the movement of a vehicle under the water-AUV. To achieve robust control of the system, two algorithms have been implemented. The first algorithm is known as Moth-Flame Optimization, inspired by the movement of moths, and another algorithm based on quantum physics called Simulated Annealing. These two algorithms are used to obtain the gain coefficients of the controllers to stabilize the steering management of the AUV. The step responses of the FOPID and LPID controller are compared. It turns out that the response obtained for the FOPID controller is superior to the LPID controller in terms of overshoot, rise time, settling time and ITAE. Furthermore, we also observe that MFO tuned FOPID controller has better results than SA tuned FOPID controller.

REFERENCES

- [1] S. K. Valluru and H. Thareja, "Evaluation of moth-flame optimization, genetic and simulated annealing tuned PID controller for steering control of autonomous underwater vehicle," in *IEEE International IOT, Electronics and Mechatronics Conference*, 2021, pp. 1–6.
- [2] B. Jalving, "The NDRE-AUV Flight Control System," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 4, pp. 497–501, 2007.
- [3] A. J. Healey and D. Lienard, "Multivariable Sliding-Mode Control for Autonomous Diving and Steering of Unmanned Underwater Vehicles," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 3, pp. 327–339, 1993.
- [4] G. V. Lakhekar, L. M. Waghmare, and P. G. Jadhav, "Robust Diving Motion Control of an Autonomous Underwater Vehicle Using Adaptive Neuro-Fuzzy Sliding Mode Technique," *IEEE Access*, vol. 8, pp. 109891–109904, 2020.
- [5] Q. Chen, T. Chen, and Y. Zhang, "Research of GA-based PID for AUV motion control," in *2009 IEEE International Conference on Mechatronics and Automation, ICMA 2009*, 2009, pp. 4446–4451.
- [6] K. H. Ang, G. Chong, and Y. Li, "PID Control System Analysis , Design , and Technology," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 559–576, 2005.
- [7] Sudarshan.K.Valluru, R. Kumar, and R. Kumar, "Design and Implementation of L-PID and IO-PID Controllers for Twin Rotor MIMO System," in *IEEE International Conference on Power Electronics, Control and Automation (ICPECA)*, 2019, pp. 1–5.
- [8] S. K. Valluru and M. Singh, "Stabilization of Nonlinear Inverted Pendulum System Using MOGA and APSO Tuned Nonlinear PID Controller," *Cogent Engineering*, vol. 4, no. 1, pp. 1–15, 2017.
- [9] S. K. Valluru, M. Singh, and S. Singh, "Prototype Design and Analysis of Controllers for One Dimensional Ball and Beam System," in *1st IEEE International Conference on Power Electronics. Intelligent Control and Energy Systems*, 2016, pp. 1–6.
- [10] C. W. Chen, J. S. Kouh, and J. F. Tsai, "Modeling and simulation of an AUV simulator with guidance system," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 2, pp. 211–225, 2013.
- [11] J. Wan, B. He, D. Wang, T. Yan, and Y. Shen, "Fractional-order PID motion control for AUV using cloud-model-based quantum genetic algorithm," *IEEE Access*, vol. 7, pp. 124828–124843, 2019.
- [12] C. Yu, A. A. Heidari, and H. Chen, "A quantum-behaved simulated annealing algorithm-based moth-flame optimization method," *Applied Mathematical Modelling*, vol. 87, pp. 1–19, 2020.
- [13] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowledge-Based Systems*, vol. 89, pp. 228–249, 2015.

Design and Realization of Wideband 4 – Element Printed MIMO Antenna for X -Ku- Band Satellite and Limited K – Band Short range Applications

Ch Murali Krishna
Research Scolar, Department of ECE,
PDPM IIIT Design and Manufacturing,
Jabalpur, India
Krishnasri780@gmail.com

Er Jeetamitra Satapathy
Govt. Industrial Training Institute,
Bhubaneswar, Odisha, India
jeetamitras@gmail.com

N Suguna
Research Scholar, School of Electronics
Engineering, Vellore Institute of
Technology, Vellore, India
reachsuguna@gmail.com

Sunita Singh
Govt. Industrial Training Institute,
Bhubaneswar, Odisha, India
principalitibbsr@gmail.com

Ganta Sandeep V Padmakar
Assistant Professor, Department of ECE,
Ramachandra College of Engineering,
Eluru, India
gsandeep4567@gmail.com

Rayudu Bharath Kumar
Assistant Professor, Department of ECE,
Vishnu Institute of Technology,
Bhimavaram, India
07ar1a0403@gmail.com

Abstract—This present works focus on increasing the isolation between antennas and enhanced radiation efficiency (η) of the proposed novel four port Multiple – Input Multiple – Output (MIMO) antenna operating in the X – band (8 - 12GHz), Ku – band (12 – 18GHz) and limited K – band (18 - 26.5GHz) applications. The design is fed by tapered structure and has been printed on FR – 4 epoxy substrate material (dielectric constant, $\epsilon_r = 4.4$, loss tangent, $\tan\delta = 0.02$, thickness, $h = 0.8mm$) having a compact size of 16.5mm x 16.5mm. Proposed MIMO structure is resonating from 7.62 to 23.04GHz at -10dB and is achieved impedance bandwidth (IBW) of 15.42GHz, bandwidth ratio (BR) of 3.02 and fractional bandwidth (FBW) of 100.58 respectively. The proposed antenna achieves improvement in isolation (S_{21}) greater than 20dB and exhibits high radiation efficiency (η) of 94.63%. Suggested antenna performance can be studied including reflection coefficient (S_{11}), voltage standing wave ratio (VSWR < 2), isolation (S_{12}), isolation between multi ports, far – field reports, current analysis, Envelope correlation coefficient (ECC < 0.015), diversity gain (DG ~ 10dB), total active reflection coefficient (TARC < 0.3), channel capacity loss (CCL < 0.2bps/Hz) and Mean effective gain (-3dB < MEG <-12dB).

Keywords—*Pot Shape, MIMO, Compact Size, Tapered fed, Impedance bandwidth, ECC, X – Ku – K – band Applications.*

I. INTRODUCTION

With enticing features including multi-band communication, fast data rate, large capacity, good resolution, and low operational energy, the wideband antenna system has captivated the current wireless world. The performance of a single-input-single-output (SISO) system, on the other hand, is hampered by multipath propagation and space obstacles. In today's communication era, increasing data speed for wireless communication applications is critical. For particular applications, it is also desired that all parts of a MIMO system have good impedance matching, identical radiation patterns, and the same polarisation. Miniaturized MIMO antenna configurations are required in wireless devices because they might achieve minimal inter-element isolation and so degrade performance [1-3]. Different diversity strategies can be used to

increase transmission link reliability by reducing multipath fading [4-5]. The receiving system's dependability, resilience, and security improve as the number of antennas with similar spectral characteristics at the receiving terminal grows; nevertheless, space is a critical consideration. In addition, when the distance between antenna components diminishes, the effect of mutual coupling grows, affecting the MIMO system's diversity performance. Several MIMO antennas have been reported in literature to address the methodologies used for improving characteristics for wireless applications [6-13]. Authors designed dual – band two element MIMO antenna for X – band and Ku - band applications on FR – 4 material within size of 24mm x 20mm [6]. Dual frequency I – shaped MIMO antenna with vertical slot introduced in ground structure resonates at 5.66GHz and 7.53GHz within compact size of 50.54mm x 21.29mm and it is having loss correlation coefficient less than 0.5 [7]. Miniaturized T – shaped four element MIMO antenna has been designed on RT / Duroid material to cover 7.8 – 16.5GHz for X – band and Ku – band applications with enhanced radiation efficiency up to 92% [8]. Two trapezoidal shaped dual – port MIMO antenna has been designed to cover C – band X – band applications within range from 6.6 – 7.6GHz and 8.3 – 10GHz and it has having high correlation between antennas less than 0.015 [9]. In [10], a four element MIMO antenna achieved circular polarization using Tai Chi – shape and L – shape structures on FR – 4 substrate designed on compact size of 46.7mm x 46.7mm. Minimizing the Inter element coupling and increasing the isolation by inserting L – shaped geometries in four element MIMO structures [11]. In [12], circularly polarized MIMO antenna is developed for X – band application over spectrum 8.07 – 11.59GHz and its isolation is better than 20dB. Two decoupling structures are used for reducing the mutual coupling between antennas. The realized impedance bandwidth of designed antenna covers 2.35GHz (1.65 – 4GHz) having fractional impedance bandwidth of 83.18% [13].

This work provides a novel design of an antenna for wideband MIMO applications with higher functionality, which

deals with mutual coupling reduction and boosting the isolation feature. A 4 – element tapered fed MIMO antenna with partial ground has been printed on FR – 4 epoxy material within a compact size of 16.5mm x 16.5mm x 0.8mm is proposed in this paper. The proposed antenna cover wide band applications which covers X – band (8 – 12GHz), Ku – band (12 – 18GHz) and limited K – band (18 – 26GHz) applications. Design methodology, simulation results analysis, advantages of suggested antenna with respect to existed works and conclusion of this work are discussed in following sections.

II. DEVELOPMENT OF PROPOSED MIMO ANTENNA EVALUATION

Figure 1 shows the three dimensional view of proposed single element and 4 – element multiple – input multiple – output (MIMO) antenna design, which are printed on two sided FR – 4 epoxy glassy substrate material with dielectric constant (ϵ_r) of 4.4, thickness (h) 0.8mm and loss tangent ($\tan\delta$) 0.02 using Ansys electromagnetic high frequency structure simulator [14]. The compact size of single element is 10.5mm x 5.3mm and MIMO antenna is 16.5mm x 16.5mm.

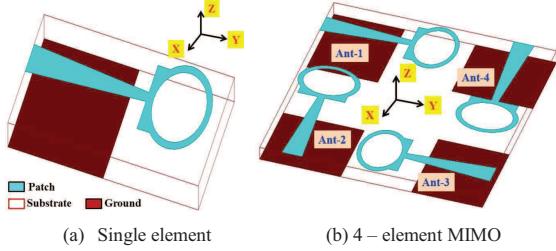


Fig. 1. Proposed Antenna design in 3D view (a) Single element (b) 4 – element MIMO Antenna

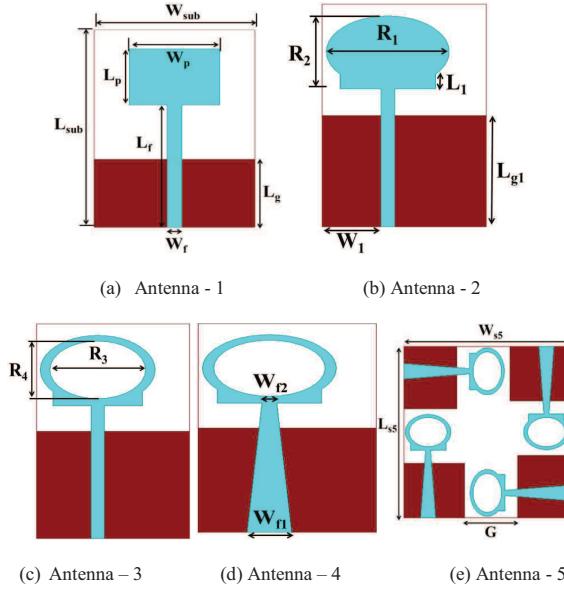


Fig. 2. Design evaluations of proposed 4 – element MIMO antenna

TABLE I. OPTIMIZED DIMENSIONS OF PROPOSED ANTENNAS ARE GIVEN IN MM [MM: MILLIMETER]

L_{sub}	10.5	L_p	3	L₁	0.7	R₃	1.9	L_{s5}	16.5
W_{sub}	5.3	W_p	3	W₁	2.1	R₄	1.4	W_{s5}	16.5
L_f	6.5	L_g	3.6	R₁	2.25	W_{n1}	1.5	G	5.25
W_f	0.5	L_{g1}	5.25	R₂	1.65	W_{n2}	0.5		

Figure 2 shows the geometrical evaluations of the proposed antenna in five steps. Conventional rectangular microstrip patch antenna using transmission line fed geometries are computed from standard mathematical expressions [15], which is represented as Antenna - 1. In step – 2. Asymmetric fed modified rectangular structure in addition with elliptical geometry, which looks like pot shape called as Antenna – 2. Elliptical slot is introduced in Antenna – 2 design, which enhances the impedance bandwidth of designed antennas names as Antenna – 3. Bandwidth (BR) can be improved by replacing line fed with tapered fed structure named as Antenna – 4. This proposed single element can be arranged in 4 – element MIMO orthogonal mode with a distance (G) of 5.25mm between antennas represented as Antenna – 5.

[1]. For an efficient radiation, the width of the microstrip patch antenna can be expresses as

$$w_p = \frac{c}{2f_r} \sqrt{\frac{2}{\epsilon_r + 1}} \quad (1)$$

Where c = free – space velocity of light = 3×10^8 m/s, f_r = resonant frequency, ϵ_r = dielectric constant

[2]. Due to fringing effect, electromagnetic waves travel some in dielectric material and some in the air. Effective dielectric constant (ϵ_{eff}) is considered to account fringing and is given by

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + 12 \frac{h}{w} \right]^{-\frac{1}{2}} \quad (2)$$

[3]. Because of fringing effects, electrical length is greater than physical dimensions. Extended electrical length of patch (ΔL) can be computed using

$$\frac{\Delta L}{h} = 0.412 \frac{\left(\epsilon_{eff} + 0.3 \right) \left(\frac{w}{h} + 0.264 \right)}{\left(\epsilon_{eff} - 0.258 \right) \left(\frac{w}{h} + 0.8 \right)} \quad (3)$$

[4]. Effective electrical length (L_{eff}) of patch antenna is sum of patch length (L) at resonant frequency and extended length (ΔL) and mathematically given as

$$L_{eff} = L + \Delta L \quad (4)$$

$$\text{Where } L = \frac{c}{2f_r \sqrt{\epsilon_r}} \quad (5)$$

[5]. The width of the transmission feed line (W_f) can write as

$$\frac{w_f}{h} = \begin{cases} \frac{8e^A}{e^{2A}-2}; & \text{if } \frac{w}{h} \leq 2 \\ \frac{2}{\pi} \left\{ B - 1 - \ln(2B-1) + \frac{\epsilon_r - 1}{2\epsilon_r} \left[\ln(B-1) + 0.39 - \frac{0.61}{\epsilon_r} \right] \right\}; & \text{if } \frac{w}{h} \geq 2 \end{cases} \quad (6)$$

Where factor A is expressed by

$$A = 2\pi \frac{Z_0}{Z_f} \sqrt{\frac{\epsilon_r + 1}{2}} + \frac{\epsilon_r - 1}{\epsilon_r + 1} \left(0.23 + \frac{0.11}{\epsilon_r} \right) \quad (7)$$

Factor B can be computed using

$$B = \frac{Z_f \pi}{2Z_0 \sqrt{\epsilon_r}} \quad (8)$$

Where Z_0 – characteristics impedance, Z_f – Wave impedance in free space

III. SIMULATION RESULTS AND DISCUSSION

In this section, simulated electromagnetic and far – field characteristics are examined. In addition to these, MIMO antenna parameters are also analysed. Figure 3 shows the return loss (S_{11}) characteristics of designed antennas. Initially rectangular microstrip patch antenna (Antenna - 1) has been designed and it is resonates at single frequency 11.20GHz with -45.97dB reflection coefficient and having impedance bandwidth of 1.91GHz (10.29 – 12.20GHz) at $S_{11} \leq -10$ dB. Rectangular patch modified to pot shape by merging elliptical structure with rectangular patch to get broad band characteristics. Antenna – 2 achieves IBW from 10.69GHz to 17.89GHz. Antenna – 3 designed (Elliptical slot has been inscribed in pot shaped patch) for better improvement in bandwidth but rejection band occurred from 12.78 to 16.13GHz because of impedance mismatching. Antenna – 4 is proposed by modifying rectangular feedline into tapered feed for impedance matching. Proposed single element obtains bandwidth of 12.98GHz covers from 9.48GHz to 22.46GHz.

Small bandwidth ratio ($BR = \frac{f_h}{f_l}$, where ' f_h ' – high resonant frequency, ' f_l ' – low resonant frequency) of the suggested antenna is 2.37. Corresponding fractional bandwidths ($FBW = \frac{f_h - f_l}{f_r} * 100$) of the simulated

antenna – 4 is 81.27%. Using the single antenna element design established in the antenna - 4, a quad element MIMO antenna (antenna - 5) is produced in this section, as shown in Figure 1. Each antenna element is $10.5 \times 5.3\text{mm}^2$ in size and is symmetrically and rotationally positioned at 90° intervals to produce a square shape. The suggested MIMO antenna achieves a port isolation of better than 20dB. It is observed that IBW based on -10dB is 15.42GHz and corresponding FBW is 100.58%. Figure 4 shows the VSWR characteristics of the proposed antenna design. Simulated characteristics of

evaluated structures are summarized in table 2. The insertion loss among the quad elements is greater than 20dB for entire bandwidth as shown in figure 5.

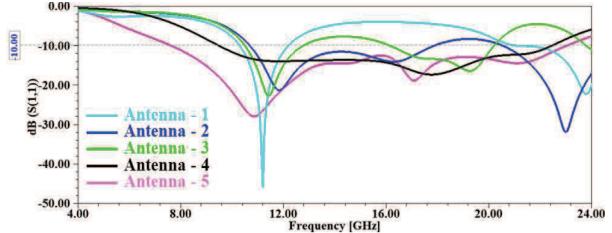


Fig. 3. Simulated reflection coefficient characteristics of design configurations

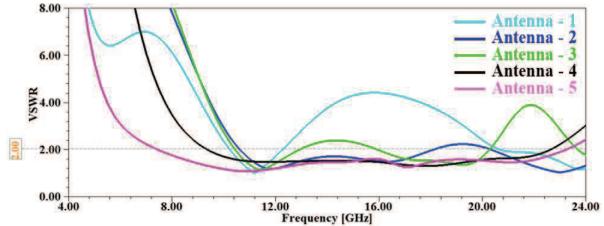


Fig. 4. Simulated VSWR characteristics of design configurations

TABLE II. CUMULATIVE DESIGN SIMULATION RESULTS OF ITERATIVE DESIGN

Design	f_l (GHz)	f_h (GHz)	IBW (GHz)	BR	FBW, %
Antenna – 1	10.29	12.20	1.91	1.18	16.98
Antenna – 2	10.69	17.89	7.20	1.67	50.38
Antenna – 3	10.49	12.78	2.29	1.21	19.68
	16.13	20.20	4.07	1.25	22.40
Antenna – 4	9.48	22.46	12.98	2.37	81.27
Antenna – 5	7.62	23.04	15.42	3.02	100.58

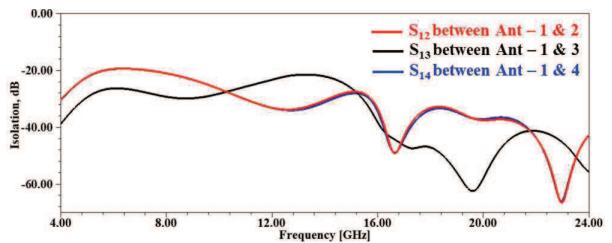


Fig. 5. Isolation characteristics between radiating elements in the designed MIMO Antenna

Figure 6 shows the three dimensional gain polar plots of proposed quad element MIMO antenna. Peak gains at

operating frequencies 10.85GHz and 17.10GHz are 3.43dBi and 4.32dBi respectively. Figure 7 represents the co – and cross – polarization patterns of the four – element MIMO antenna at 10.85GHz and 17.10GHz in E – plane (YZ – plane) and H – plane (XZ – plane) respectively.

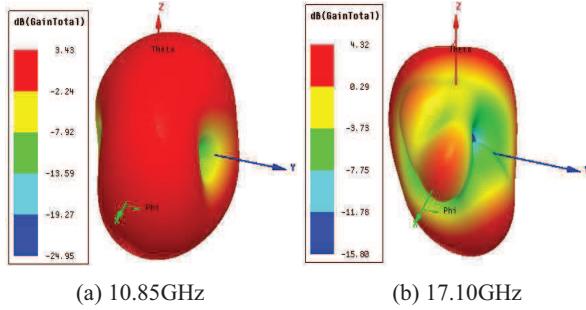


Fig. 6. Three dimensional gain polar plots at resonant frequencies

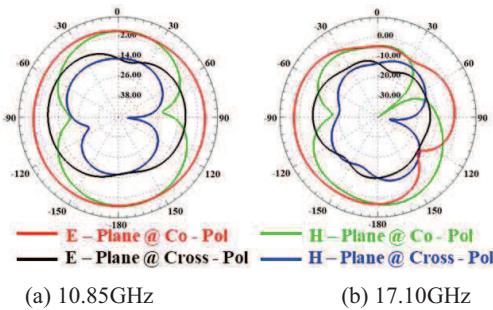


Fig. 7. Two dimensional radiation patterns at resonant frequencies due to excitation of E – plane and H – plane interms of Co – and Cross – polarizations

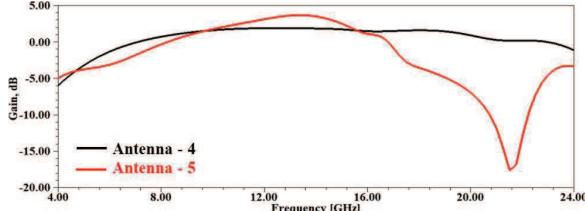


Fig. 8. Gain plot of single element and 4 – element MIMO antenna

Figure 8 shows the antenna gain characteristics of the MIMO antenna designed on FR - 4 material. It can be noticed that peak gain of single element and 4 – element MIMO antenna are 1.86dB and 3.65dB respectively. Radiation efficiency (η) characteristics of proposed MIMO antenna are given in figure 9. Simulated radiation efficiencies of single element and proposed MIMO antenna are 92.34% and 94.63% respectively. However radiation efficiency can also be compared in section 4. It can be analysed that from table 3, radiation efficiency compared with the existed works gives

better performance. The simulated surface current distribution of proposed antenna at 10.85GHz and 17.10GHz frequencies are shown in figure 10.

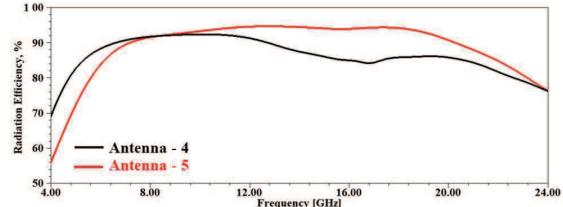


Fig. 9. Radiation efficiency representation of single element and 4 – element MIMO antenna

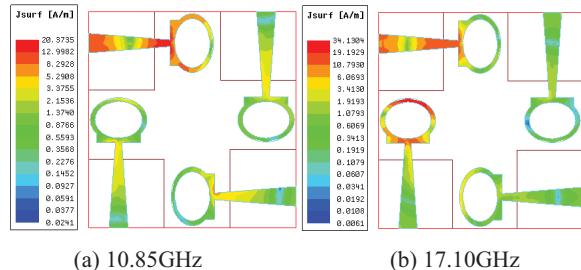


Fig. 10. Current distribution at resonant frequencies of proposed MIMO

MIMO antenna performance is investigated in terms of ECC, DG, TARC, CCL and MEG, which are computed using standard mathematical expressions derived from S – parameters [4-5, 16 – 18]. Figure 11 – 15 shows the MIMO antenna characteristics ECC, DG, TARC, DG and MEG of quad elements respectively. For uncorrelated MIMO antennas, the optimum value of ECC is 0, however 0.5 is a reasonable number in practise [16]. Figure 11 illustrates the MIMO antennas' error correlation coefficient (ECC) over the operational spectrum, which is less than 0.015. Figure 12 shows a MIMO-designed diversity gain (DG) of approximately 10 dB over the entire frequency range. Figure 13 shows the TARC plot of quad element MIMO antennas. Figure 14 represents the Channel capacity loss (CCL) of 4 – element designed MIMO antenna. The proposed antenna obtains acceptable values (ideally CCL values should be less than 0.4 bits/S/Hz within the operating band) [18]. MEG of individual elements in quad element MIMO are reported in figure 15, which is in the acceptable limits over the operating band.

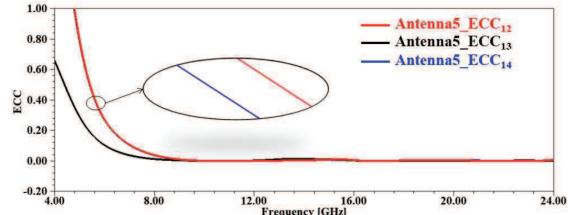


Fig. 11. Envelope correlation coefficient characteristics of suggested 4 – element MIMO antenna

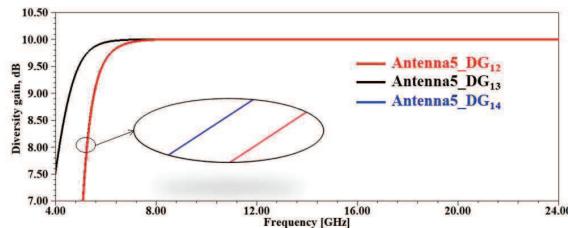


Fig. 12. Diversity gain characteristics of suggested 4 – element MIMO antenna

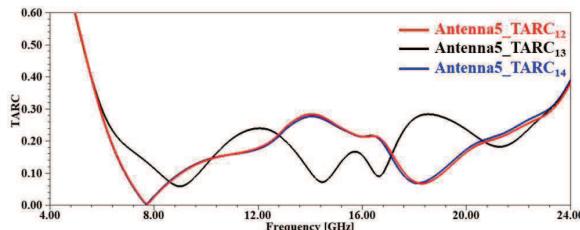


Fig. 13. TARC characteristics of suggested 4 – element MIMO antenna

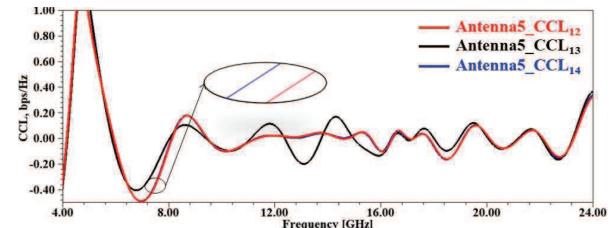


Fig. 14. Channel capacity loss characteristics of suggested 4 – element MIMO antenna

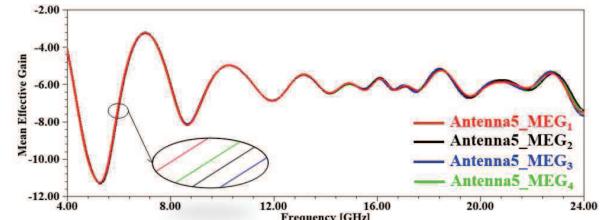


Fig. 15. Mean effective gain characteristics of suggested 4 – element MIMO antenna

TABLE III. COMPARISON OF PROPOSED 4 – ELEMENT MIMO ANTENNA WITH EARLIER WORKS

Ref.	N	Dimension, mm ²	Frequency band, GHz	BW, GHz	BR	FBW, %	Isolation, dB	ECC	η , %	Used Material
[6]	2	24 x 20	7.6 – 8.16	0.56	1.07	7.10	> 15	< 0.04	80	FR - 4
			13.8 – 14.4	0.6	1.04	4.25				
[7]	2	50.54 x 21.29	5.55 – 5.86	0.31	1.05	5.43	> 15	< 0.5	NR	FR - 4
			7.93 – 8.36	0.43	1.05	5.27				
[8]	4	25 x 25	7.8 – 16.5	8.7	2.11	71.60	> 15	< 0.14	92	RT Duroid
[9]	2	17 x 42	6.6 – 7.6	1	1.15	14.08	> 17	< 0.015	70 - 80	FR - 4
			8.3 – 10	1.7	1.20	18.57				
[10]	4	46.7 x 46.7	7.58 – 8.04	0.46	1.06	5.9	> 20	< 0.003	68	FR - 4
			9.23 – 10.79	1.56	1.17	15.6				
[11]	4	40 x 40	2.70 – 4.94	2.24	1.83	58.6	> 11	< 0.1	NR	FR - 4
[12]	2	29 x 23	8.07 – 11.59	3.52	1.43	35.80	> 20	< 0.01	85 - 93	FR - 4
[13]	2	70 x 120	1.65 – 4	2.35	2.42	83.18	> 20	NR	NR	Rogers RO3203
Proposed	4	16.5 x 16.5	7.62 – 23.04	15.42	3.02	100.58	> 20	< 0.015	94.63	FR - 4

Notation: N – Number of elements, BW – Bandwidth, BR – bandwidth ratio, FBW – Fractional bandwidth, ECC – Envelope correlation coefficient, η – radiation efficiency, NR – Not reported

IV. COMPARISON OF PROPOSED WORK WITH EXISTED WORKS

In this section, proposed antenna advantages are presented in table 3. Details of some electromagnetic characteristics and MIMO antenna parameters and some existed MIMO antennas are tabulated in table 3. It is observed that proposed antenna has huge IBW compared to earlier works, good isolation, better BR, high FBW, low correlation between ports, high radiation efficiency and miniaturized in size. According to the parameters reported in table 3, proposed antenna exhibits innovative and novel design.

V. CONCLUSION

A wideband four - element MIMO antenna with tapered fed is proposed. The main objective of this work is to enhance impedance bandwidth, reducing isolation between antennas and increasing radiation efficiency. MIMO antenna performances are investigated interms: i) ECC is less than 0.015 which is below the critical value of 0.5 ii) DG is nearly equals to 10dB iii) TARC is less than 0.3 iv) CCL is less than 0.2bits/sec/Hz v) MEG is in between -3dB and -12dB. It is noticed that four elements are placed in orthogonal orientation leads to improvement in IBW up to 15.42GHz with peak radiation efficiency of 94.63%, which states that it is suitable for X – band (8 – 12GHz) & Ku – band (12 – 18GHz) for satellite based applications and K – band (18 – 26.5GHz) for short range wireless applications.

REFERENCES

- [1] W. M. Saadh, K. Ashwath, P. Ramaswamy, T. Ali, and J. Anguera, "A uniquely shaped MIMO antenna on FR4 material to enhance isolation and bandwidth for wireless applications," AEU-International Journal of Electronics and Communications, vol. 123, Article ID 153316, 2020.
- [2] Xia, X.-X., Q.-X. Chu, and J.-F. Li, "Design of a compact wideband MIMO antenna for mobile terminals", Progress In Electromagnetics Research C, Vol. 41, 163-174, 2013.
- [3] Debdeep, S. and V. K. Srivastava, "A compact four-element MIMO/diversity antenna with enhanced bandwidth", IEEE Antennas and Wireless Propagation Letters, Vol. 16, 2469-2472, 2017.
- [4] Ch Murali Krishna, M. Sai Prapoorna, K. Taruni Sesha Sai, M. Sai Teja, "Super Wideband 1×2 MIMO Antenna for Advanced Wireless Communication", Springer Advances in Electrical and Computer Technologies, <https://doi.org/10.1007/978-981-15-9019-1>
- [5] C. M. Krishna, M. Mamatha and J. P. Kumar, "Mickey Mouse Modeled MIMO Antenna for Extended UWB Applications," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 522-527, doi: 10.1109/ICCMC48092.2020.ICCMC-00097.
- [6] El Ouahabi, M., A. Zakriti, M. Essaaidi, A. Dkiouak, and E. Hanae, 'A miniaturized dual-band MIMO antenna with low mutual coupling for wireless applications," Progress In Electromagnetics Research, Vol. 93, 93 - 101, 2019.
- [7] Satam, V., & Nema, S. (2015). Defected ground structure planar dual element MIMO antenna for wireless and short range RADAR application. 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES). doi:10.1109/spices.2015.7091441
- [8] Nirdosh, C. M. Tan and M. R. Tripathy, "A miniaturized T-shaped MIMO antenna for X-band and Ku-band applications with enhanced radiation efficiency," 2018 27th Wireless and Optical Communication Conference (WOCC), 2018, pp. 1-5, doi: 10.1109/WOCC.2018.8372695.
- [9] Pouyanfar, N., C. Ghobadi, J. Nourinia, K. Pedram, and M. Majidzadeh, 'A compact multi- band MIMO antenna with high isolation for C and X bands using defected ground structure," Radioengineering, Vol. 27, No. 3, 686{693, 2018.
- [10] Eslami A, Nourinia J, Ghobadi C, Shokri M (2021). Four-element MIMO antenna for X-band applications. International Journal of Microwave and Wireless Technologies 1–8. <https://doi.org/10.1017/S1759078720001440>
- [11] D. Sarkar and K. V. Srivastava, "A Compact Four-Element MIMO/Diversity Antenna With Enhanced Bandwidth," in IEEE Antennas and Wireless Propagation Letters, vol. 16, pp. 2469-2472, 2017, doi: 10.1109/LAWP.2017.2724439.
- [12] Chaudhary, P., Kumar, A., & Kanaujia, B. (2020). A low-profile wideband circularly polarized MIMO antenna with pattern and polarization diversity. International Journal of Microwave and Wireless Technologies, 12(4), 316-322. doi:10.1017/S175907871900134X
- [13] Isaac, A. A., Al-Rizzo, H., Yahya, S., Al-Wahhamy, A., & Abushamleh, S. (2018). Decoupling of Two Closely-Spaced Planar Monopole Antennas Using Two Novel Printed-Circuit Structures. Microwave and Optical Technology Letters, 60(12), 2954–2963. doi:10.1002/mop.31405
- [14] ANSYS HFSS software. <http://www.ansoft.com/products/hf/hfss/>
- [15] C.A. Balanis, Antenna Theory Analysis and Design, 2nd ed., John-Wiley and SonsPublications, 1997, ISBN 978-81-265-1393-2.
- [16] R Chandel, A K Gautam and K Rambabu, "Design and packing of an eye – shaped Multiple Input – Multiple- Output Antenna with High Isolation for Wireless UWB Applications", IEEE Transactions on Comp. Pack. Manufact. Technnology., Vol 8, Issue 4, Page No 635 - 642, 2018.
- [17] Sharawi MS. Printed MIMO Antenna Engineering. Norwood, MA: Artech House; 2014.
- [18] Ojaroudiparchin N, Shen M, Zhang S, Pedersen G.F, "A switchable 3D coverage – phased array antenna package for 5G mobile terminals", IEEE Antennas Wireless Propagation Letters, 2016, 15, 1747 – 1750.

Artifical Intelligence based Real Time Lie Detector Using Eye Gaze Pattern

¹Krishna Kishore Koganti

Dept. of ECE

GMR Institute of Technology

Rajam,India

krishnakishore.k@gmrit.edu.in

²SSD Kameswari

Dept. of ECE

GMR Institute of Technology

Rajam,

Durgakameswari.s@gmrit.edu.in

³Dr V Jagan Naveen

Department of ECE

GMR Institute of Technology

Rajam,

Jagannaveen.v@gmrit.edu.in

Abstract— Deception is a sophisticated talent that takes place in human social interactions. Lying may be described as the act of hiding the truth using a false statement. Traditional methods like polygraphs, which are used for lie detection, involve the monitoring of sweat, respiratory rate, pulse rate, and blood pressure. These methods are not only invasive but also expensive. However, it also requires a skilled human interviewer to perform the questioning and evaluate the data. Moreover, those techniques are not always reliable. The need for an improvised way of detecting the deceitfulness of a person has become critical in the society we live in. The lie detection here is based on the hypothesis that liars experience more cognitive demand than truth-tellers. Lie detection will be performed in real-time video by capturing the eye gaze pattern. The gaze can be either to the left or to the right, which is based on the truthfulness of the person in the video. The eye gaze pattern is analyzed with the help of OpenCV using the dlib library and a pre-trained network called shape_predictor_68_landmarks.

Keywords— Deception, Polygraph, Invasive, Cognitive demand, Eye gaze, OpenCV, Dlib.

I. INTRODUCTION

Humans are frequent liars, and such frequent lying should make us expert liars, but it is not the case. Several research suggestions suggest that lying is more cognitively demanding than truth-telling. For instance, while lying, instead of simply holding one story in memory (the truth), two stories are held in memory; the truth and the lie. Additionally, the liar must hold both of the stories separate to avoid the truth contaminating the lie. Still, the liar must make sure the details of the lie match peripheral details surrounding the story [1]. Furthermore, one must also be aware of and appropriately manage verbal and non-verbal indicators that might suggest deceptive behavior [2]. Therefore, poor management of behavior relevant to telling a convincing lie strains cognitive ability and reduces the amount of control a liar has when telling a lie [3]. The result of such poor management and lack of control can trigger involuntary actions such as pupil dilation, high blink rate, increased heartbeat, eye gaze in different directions etc. Some of the traditional methods of lie detection based on these cues are polygraph and CVSA, These techniques are invasive and expensive [4]. Moreover, these are not always reliable and require a professionally trained person to interpret the deceit from the results. The main aim of the project is to design a lie detector, which should be simple and there should be no need for any interpretation of the results. This is done by designing a lie-detection algorithm that analyses the behavioral cue, which is an eye gaze pattern

caused by excessive cognitive demand. The output is based on the notion that for a right-handed person, the gaze is directed towards the right when he is remembering something and to the left while he is constructing something that he hasn't experienced. The entire project is designed using OpenCV in PyCharm IDE with a Python interpreter so that the lie detection can be performed using both a desktop with a webcam and a Raspberry Pi with a connected camera.

II. LITERATURE SURVEY

Yulia V. Bessonova and Alexander A. Oboznov [5] presented the paper based on the important technique that is the polygraph testing technique in comparison with the eye moments. This testing technique is a widespread method, and the solutions it provides are not very difficult to achieve. Most of the arguments included in it are based on the low validity of polygraph testing and the high variability of individual responses. Jeffrey J. Walczyk [6] suggested the study based on Time Restricted Integrity-Confirmation. In the testing procedure, they tested it by telling "witnesses" to genuine crime recordings to lie or tell the truth to relevant questions. He presented a revolutionary strategy for lie detection aimed at increasing the cognitive burden on liars and decreasing it on truth-tellers by preventing the recital of dishonest responses. B. Singh, P. Rajiv, and M. Chandra [7] suggested a method for lie detection and identifying the truth by using image processing techniques. The lie detection mechanism is based upon the scientifically verified concept obtained from eye blink literature that liars incur higher cognitive demands than truth-tellers. Thus, their falsehoods are accompanied by a large drop in eye blinks, closely followed by a rise in eye blinks when the cognitive load relaxes, once the lie has been uttered. Blink detection is done with the aid of MATLAB, utilizing the HAAR Cascade method. R. H. Nugroho, M. Nasrun, and C. Setianingsih [8] suggested a Lie Detector with Pupil Dilatation and Eye Blinks using Hough Transform and the Frame Difference Method with Fuzzy Logic. Looking at pupils can help you determine whether someone is lying or not because, according to psychology, pupils dilate when a person is in a state of despair, even if they are lying. Eye blinking may also be an indicator to recognize those who lie or not by increasing the amount of blinking. Accuracy is gained through observations of the change in pupil diameter using the circular Hough transform technique and increasing the number of blinks of the eye using the frame difference method.

III. METHODOLOGY

Based upon the direction of eye gaze, it can be revealed whether the person is making a truthful statement or not [9]. Different eye gaze directions are shown in figure 3.1.

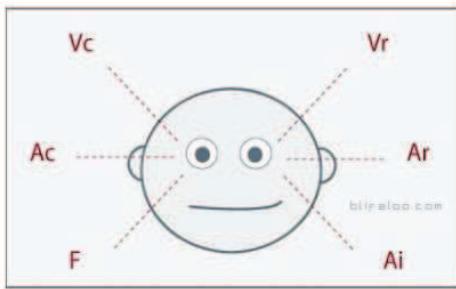


Figure 3.1: Eye gaze directions

- Ac: Auditory Constructed
- Ar: Auditory Remembered
- Vc: Visually Constructed
- Vr: Visually Remembered
- F: Feeling / Kinesthetic
- Ai: Internal Dialog

A. Eye gaze pattern

For tracking the eye gaze, real time video of a person is used. After getting the live video, Face should be identified, using the futures in the live video and, after that, the eyes need to be located on the face [10]. For finding the face from live video and the eyes from a pre-trained network called "shape_predictor_68_landmarks" is used. This pre-trained network finds the faces in the video with 68 landmark points, as shown in Figure 3.6.

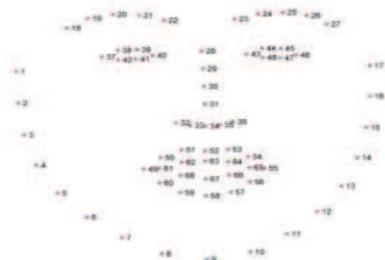


Figure 3.2: Face detection with 68 landmarks

After finding the face with 68 landmarks, the eyes are separated into a separate frame by using the landmark points of the eyes. The landmarks of different parts of the face are mentioned in table 3.1.

Table 3.1: Facial Landmarks

Facial part	Landmarks
Right eye	(37,38,39,40,41,42)
Left eye	(43,44,45,46,47,48)
Nose	(28,29,30,31,32,
Mouth	(49 to 68)

B. Block diagram

The real-time video will be accessed through the webcam on the desktop. This video will be analysed for eyeball gaze,

and the number of white pixels in the eye frames will be given as input to the lie-detecting algorithm. The algorithm processes these inputs to give an output at the end.

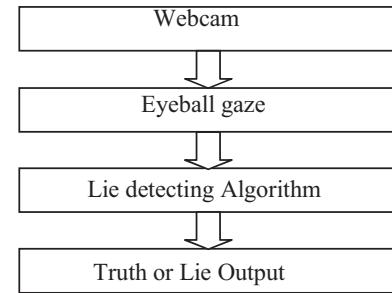


Figure 3.3: block diagram

C. Implementation process

Lie detection by tracking the eye gaze is not a single step process. It takes several steps to do it. The very first step is to get the live video from the webcam. And this real time video is processed through the following steps to give the output. The step-by-step process is shown figure 3.3.

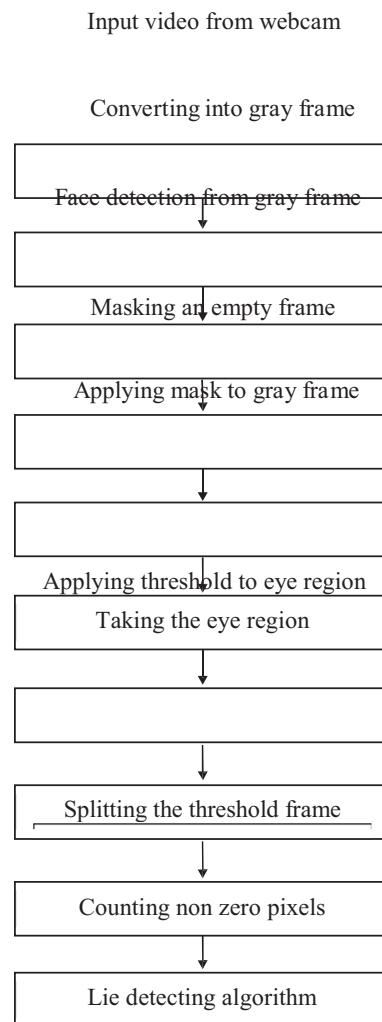


Figure 3.4: Sequence of Lie detection

D. Step 1

As the lie detection should be done on live video, first aim is to get access to the live video as shown in figure 3.5, which is done using OpenCV. The real-time video can be accessed by using the cv2.VideoCapture() command. The video capture can be terminated by using the OpenCV command cv2.waitKey().



Figure 3.5: Live video frame

Step 2

After getting access to live video, eyes have to be located in order to track the gaze. For this purpose, a pre-trained network called shape_predictor_68_landmarks is used. This pre-trained network is used with the help of the dlib library. All the commands of the dlib library work on the gray frame only. So, after getting access to the video, the next step is to convert the live video frame into a gray frame as shown in figure 3.6. cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY) is the OpenCV command used to convert to a gray frame.



Figure 3.6: Gray video frame

Step 3

After converting the live video frame into a gray frame, the next step is to detect the face from the gray frame, as shown in figure 3.7. This can be done with the help of a pre-trained network "shape predictor 68 landmarks" which detects faces with 68 key-points. For this purpose, we use an inbuilt function in the dlib library called detector. Face detection is performed using the commands detector=dlib.get_frontal_face_detector() and faces=detector(frame_name).

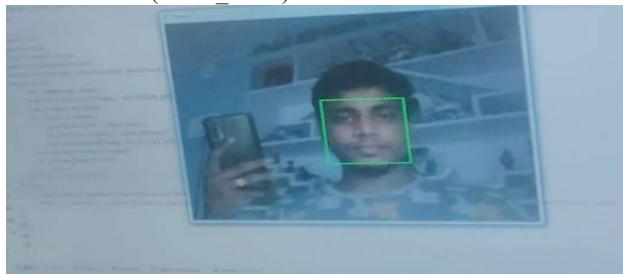


Figure 3.7: Face detection

Step 4

After detecting the face, in order to get the eye region from the gray frame, it needed to be masked, leaving only the eye region as shown in figure 3.8. This can be done by taking an empty frame of the same dimensions as that of the gray frame, masking it with black, and clearing the eye region with the help of landmarks. The OpenCV command used for masking an entire frame is mask = np.zeros (8 height, width), np.uint8) & cv2.fillPoly (mask, [left_eye_region], 255), where u, 8 in unit8, means unsigned and the number of bytes of each pixel, respectively. The number 255 indicates the maximum value in gray scale and height, while width is the dimensions of the gray frame.

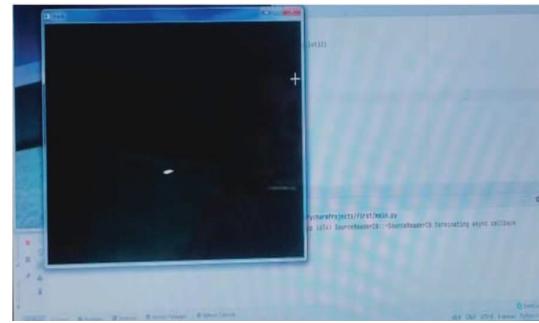


Figure 3.8: Mask frame

Step 5

After getting the masked frame, this masked frame must be applied to the gray frame in order to separate just the eye region from the gray frame, as shown in figure 3.8. Masking helps us focus on the eye portion. The OpenCV command used for masking is new_frame_name = v2.bitwise_and (gray, gray, mask=mask_name). Here, each pixel value in the gray frame is recalculated based on the mask



Figure 3.9: Masked gray frame

Step 6

Now, after masking the gray frame, with the help of key-points of the eye, the eye region is to be separated to put it in a new frame from the masked gray frame. The new frame will be as shown in figure 3.10. The OpenCV command used for this purpose is new_frame = masked_gray_frame [min_y: max_y, min_x: max_x]. Here, min_y, max_y, and min_x, max_x are the minimum and maximum key-points of the eye region.



Figure 3.10 Eye region from masked gray frame

Step 7

After concentrating on the eye region with the aid of masking, the method of thresholding has to be applied for this frame. Thresholding is a method in OpenCV that involves the assignment of pixel values in accordance with the threshold value specified. In thresholding, each pixel value is compared with the threshold value. If the pixel value is lower than the threshold, it is set to 0, otherwise it is set to the maximum value (255). (255). After thresholding, the eye will be as illustrated in picture 3.10. The OpenCV command used for this operation is new frame = cv2.threshold (gray eye, thresholdvalue, 255, cv2.THRESH_BINARY).



Figure 3.11 Threshold eye

Step 8

Now the threshold eye frame is to be accessed as two separate half frames as shown in the figure.3.12, as well as Figure 3.13. This is done to calculate the number of white pixels in each frame. The threshold eye frame is separated in a way such that if the dimensions of the threshold eye frame are "0 to w" along the x-axis and "0 to h" along the y-axis, then the dimensions of the left side frame will be "0 to w/2" along the x-axis and "0 to h" along the y-axis, and the dimensions of the right side frame will be "w/2 to w" along the x-axis and "0 to h" along y-axis.

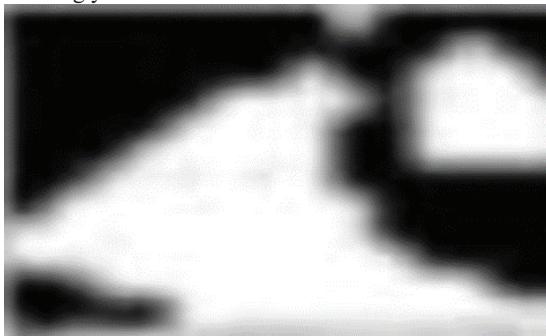


Figure 3.12: Left half of threshold eye

The above image is the left-side half of the threshold eye. If the person looks to his left, the number of white pixels in the above frame will decrease, and if he looks to his right, the number of white pixels in the above frame will increase.



Figure 3.13: Right half of threshold eye

The above image is the right-side half of the threshold eye. If the person looks to his left, the number of white pixels in the above frame will increase, and if he looks to his right, the number of white pixels in the above frame will decrease. These values from both the frames will act as inputs to the lie-detecting algorithm.

IV Lie Detector Algorithm

The nonzero pixel values from both half frames of the threshold image are the inputs to the lie-detecting algorithm. The input to the lie-detecting algorithm consists of two values as shown in figure 4.1.



Figure 4.1: Inputs to lie detecting algorithm

The lie detecting algorithm works in the following way

Step 1:

Lie detecting algorithm takes two input values.

Step 2

The two values will be divided to get the gaze ratio. (Upon repeatedly executing the process by looking to the left and right and analysing the gaze ratio, the gaze ratio seems to vary from 0.5 to 2.8. From this range, two points of threshold are selected and coded into the algorithm.

Step 3

When the gaze ratio is less than the lower threshold value, the gaze direction will be assigned as left. Similarly, when the gaze ratio is greater than the higher threshold value, the gaze direction will be assigned as right, and when the gaze ratio is between the lower and higher threshold, the gaze will be assigned as centre by the lie-detecting algorithm.

Step 4

The lie detecting algorithm keeps track of these gaze directions throughout the video. And at the end of the video, the algorithm counts the number of gaze directions.

Step 5

If the gaze is more towards the left, then according to cognitive load theory, the person should be constructing something that he hasn't experienced, and similarly, if the gaze is more towards the right, then the person should be remembering something. According to this notion, the algorithm scores whether the person is saying the truth or a lie. The sequence of steps followed by the lie detecting algorithm is shown in Figure 4.2.

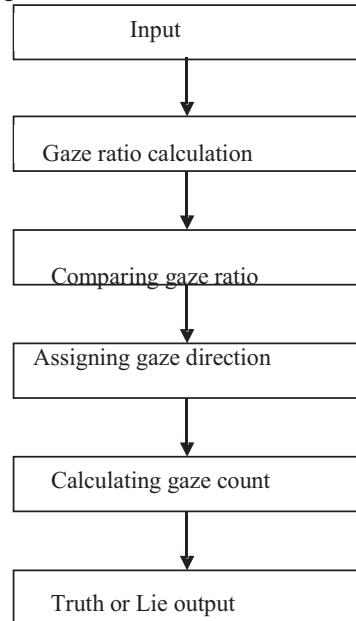


Figure 4.2: Sequence of lie detecting algorithm

V RESULTS

The simulation results at each step while tracking the eye gaze are already demonstrated in chapter 3. The final findings are displayed in figure 5.1, figure 5.2 and figure 5.3. With three use cases.

Case 1

In the first case the individual is requested to state the truth and the results were evaluated, apart from situations which rendered the person distracted due to external disturbances, the output of lie detector is correct most of the times.



Figure 5.1: Person taking lie detection test

```

# memory
frame = cv2.imread("frame.jpg", cv2.IMREAD_COLOR)
# cv2.lineDetector(frame, (left_eye_region), True, (0, 0, 255), 2)
height, width, depth = frame.shape
mask = np.zeros((height, width), np.uint8)
cv2.lineDetector(frame, (left_eye_region), True, 255, 2)
cv2.lineDetector(frame, (left_eye_region), 255)
cv2.lineDetector(mask, (left_eye_region), gray, mask=mask)
cv2.imshow("eye", eye)
min_x = np.min(left_eye_region[:, 0])
max_x = np.max(left_eye_region[:, 0])
min_y = np.min(left_eye_region[:, 1])
max_y = np.max(left_eye_region[:, 1])
gray_eye = np.array([min_x, min_y, max_x, max_y])
threshold_eye = cv2.THRESH_BINARY
height, width = threshold_eye.shape
get_rect_val()

```

Figure 5.2: Truth output

```

# memory
frame = cv2.imread("frame.jpg", cv2.IMREAD_COLOR)
height, width, depth = frame.shape
mask = np.zeros((height, width), np.uint8)
cv2.lineDetector(frame, (left_eye_region), True, 255, 2)
cv2.lineDetector(frame, (left_eye_region), 255)
cv2.lineDetector(mask, (left_eye_region), gray, mask=mask)
cv2.imshow("eye", eye)
min_x = np.min(left_eye_region[:, 0])
max_x = np.max(left_eye_region[:, 0])
min_y = np.min(left_eye_region[:, 1])
max_y = np.max(left_eye_region[:, 1])
gray_eye = np.array([min_x, min_y, max_x, max_y])
threshold_eye = cv2.THRESH_BINARY
height, width = threshold_eye.shape
get_rect_val()

```

Figure 5.3: Lie output

Case 3

When the number of times the person is looking to the left is equal to number of times looking to the right, then it shows the output as "Try again" as shown in figure 5.4.

```

# memory
frame = cv2.imread("frame.jpg", cv2.IMREAD_COLOR)
gaze_ratio_left_eye = get_gaze_ratio([17, 21, 39, 40, 41], landmarks)
gaze_ratio_right_eye = get_gaze_ratio([42, 43, 44, 45, 46, 47], landmarks)
gaze_ratio_left_eye, gaze_ratio_right_eye, ratio, gaze_ratio_right_eye - ratio
if gaze_ratio_right_eye > ratio:
    L.append("RIGHT")
    cv2.putText(frame, "RIGHT", (50, 100), font, 2, (0, 0, 255), 3)
    cv2.putText(frame, "TRY AGAIN", (10, 100), font, 1, (0, 0, 255), 1)
    cv2.putText(frame, "LEFT", (10, 100), font, 1, (0, 0, 255), 1)
    cv2.putText(frame, "CENTRE", (10, 100), font, 1, (0, 0, 255), 1)
    cv2.imshow("frame", frame)
    cv2.waitKey(1000)
    cv2.destroyAllWindows()
else:
    L.append("LEFT")
    cv2.putText(frame, "LEFT", (50, 100), font, 2, (0, 0, 255), 3)
    cv2.putText(frame, "TRY AGAIN", (10, 100), font, 1, (0, 0, 255), 1)
    cv2.putText(frame, "CENTRE", (10, 100), font, 1, (0, 0, 255), 1)
    cv2.imshow("frame", frame)
    cv2.waitKey(1000)
    cv2.destroyAllWindows()
    cv2.imshow("gray", gray)
    cv2.waitKey(1000)
    cv2.destroyAllWindows()

```

Figure 5.4: Try again output

VI CONCLUSION

In this a novel approach is used for efficiently tracking the eye gaze pattern of a person from a real-time video. The eye gaze pattern acquired as a consequence of successful implementation of this approach can be utilized for checking if a person is telling the truth or not. This procedure is empirically confirmed to function by testing on different folks and the outcomes are good in most of the situations. In order to perform this procedure with great efficiency strong light conditions are essential. With bad light conditions the eyes will not be fully identified and lie detection

becomes inaccurate. The surrounding settings of the individual being tested must be peaceful and should not be bothersome while the video is being evaluated for eye gazing. To be exact there should not be any scenarios that make the individuals eyes to stare in other directions. This procedure does not need any technically skilled specialist to interpret the findings. As it is based on the Scio motive behavior of the person, it can predict whether the person is lying or not.

REFERENCES

- [1] Bhaskar, Anand, Selvam Subramani, and Rajdeep Ojha. "Respiratory belt transducer constructed using a singing greeting card beeper." *Advances in physiology education* 37.1 (2013): 117-118.
- [2] John Synnott, David Dietzel & Maria Ioannou (2015) A review of the polygraph: history, methodology and current status, *Crime Psychology Review*, 1:1, 59-83.
- [3] Abboud, B., Wadkins, T.A., Forrest, K.D., Lange, J. and Alavi, S., 2002. False confessions: Is the gender of the interrogator a determining factor. In biennial meeting of the American Psychology-Law Society, Austin, TX.
- [4] Meissner, C.A., Russano, M.B. (2003). The psychology of interrogations and false confessions: Research and recommendations. *Canadian Journal of Police & Security Services*, 1, 53–64.
- [5] Yulia V. Bessonova and Alexander A. Oboznov "Eye Movements and Lie Detection" IHSI ,pp.149-155 ,2018.
- [6] Jeffrey J. Walczyk "LIE Detection by Inducing Cognitive load:Eye movements and other cues to false answers" IACPFP, vol. 39,pp. 887-909,2012.
- [7] B. Singh, P. Rajiv and M. Chandra, "Lie detection using image processing," 2015, International Conference on Advanced Computing and Communication Systems, 2015, pp. 1-5, doi: 10.1109/ICACCS.2015.7324092.
- [8] R. H. Nugroho, M. Nasrun and C. Setianingsih, "Lie detector with pupil dilation and eye blinks using hough transform and frame difference method with fuzzy logic," 2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC), 2017, pp. 40-45, doi: 10.1109/ICCREC.2017.8226697.
- [9] S. Y. Han, I. Hwang, S. H. Lee and N. I. Cho, "Gaze estimation using 3-D eyeball model and eyelid shapes," 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016, pp. 1-4, doi: 10.1109/APSIPA.2016.7820784.
- [10] M. Kassner, W. Patera and A. Bulling, "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction", Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pp. 1151-1160, 2014.

EB Algorithm for Effective Privacy and Security of Data Processing in MCC

Dr.V.S.Prakash
Department of CSE
Kristu Jayanti College
Bengaluru, India
vspakash@kristujayanti.com

Dr.R.Thiagarajan
Department of IT
Prathyusha Engineering College
Chennai, India
rthiyagarajantpt@gmail.com

Dr.N.Bharathiraja
Department of Biomedical
Veltech Multitech Dr.Rangarajan
Dr.Sakunthala Engineering College
Chennai, India
bharathinatarajan78@gmail.com

Dr.R.Krishnamoorthy
Department of ECE
Prathyusha Engineering College
Chennai, India
krishnamoorthy.mkr@gmail.com

R.Deiva Nayagam
Department of ECE
Ramco Institute of Technology
Rajapalayam, India
deivam90@gmail.com

J.Omana
Department of IT
Prathyusha Engineering College
Chennai, India
omana.it@prathyusha.edu.in

Abstract - Mobile Cloud computing (MCC) distributes amenities over the Internet, such as hardware, software, and bandwidth. Mobile gadgets smartphones are enabled for exploration. Customers and businesses are increasingly adopting mobile cloud computing technology. Network security, data access, authentication, web application security, authorization, data breach, and data confidentiality are the entire anxiety of MCC's protection. Mobile devices' data storage capacity is limited due to adequate data storage and processing. Security threats must be investigated and analyzed in order to build a secure MCC environment. The study's purpose is to identify a way out that can improve technological requirements in mobile cloud computing in terms of data privacy and security for clients. The enhanced blowfish (EB) approach is utilized to encrypt a person's data security, and the global private key is hashed using a basic hash function. The integrity and privacy of user data can be improved through hashing. The suggested technique is compared to a standard blowfish algorithm and DES with various parameters. The EB algorithm outperforms both the standard blowfish algorithm and the DES algorithm.

Keywords— Mobile Cloud computing, hash function, blowfish algorithm, DES.

I. INTRODUCTION

For example, smartphones give consumers better connectivity and access to services and applications today [1-4]. Even though mobile technology advances, modern mobile terminals are limited by processing capabilities, memory space, and disc storage [5]. By combining current computing technology, cloud computing provides a stable solution to service delivery. The majority of cloud computing implementations appear to be based on three service delivery models.

Cloud services are available to consumers in a variety of formats. Wireless cloud computing is one of them. MCC is used in connection with cloud computing because of its core qualities such as general range, elastic, on-demand, group of resources, and pay per user. Fig. 1 depicts the main principle of MCC, where the customer transfers information to a remote user's infrastructure for on-demand cloud storage, after which the user loses physical confidentiality of data. Furthermore, because the user is ignorant of their data location, there is a greater chance that a hostile user may acquire access to it and misuse it [6-9].

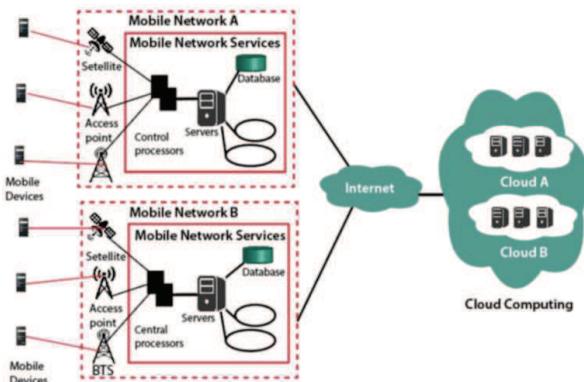


Fig.1. Example - Model for Mobile Cloud Computing

Several finite resources in consumer mobile phones, mobility management, availability shortage of channel bandwidth, stability, network access charges, elasticity, trust, application services issues, process offloading, and so on are a quantity of MCC challenges. Some of the difficulties with the evolution of MCC include the security for mobile phone users and loss of data confidentiality which are significant hurdles to the rapid taking on of MCC. Even though it is a broad subject, researchers are chosen to concentrate on its most important problems, "Complete absence of Security and privacy in MCC" [10-14]. Furthermore, to solve these challenges, we will focus on the current ways for securing user data in MCC.

A. Data Security in Mobile Cloud

Throughout its lifecycle, data protection alludes to protecting data from unauthorized users and corruption. Data encryption, tokenization, and essential management distributions are all part of this. If a company needs extra security and confidentiality, it can create a distributed system among numerous servers. Mirror Servers are a sort of server that does exactly that. Because information about an organization can be shared, data security is jeopardized. Therefore, we need data protection in MCC systems to secure such sensitive information [15-17]. The following are the essential abilities that a cloud provider must provide: a tried and true cryptography scheme for shared storage to keep data safe, very rigorous user access protocols to ensure that no

unauthorized users gain access to data, data backup must be scheduled, and media backup must be safely stored.

Security is usually within these minimum capabilities, but in attendance are at a halt some difficulties that could be addressed [18-24]. Is information safety solely the duty of the cloud provider, or is it also the responsibility of the business that leases the assets? To preserve and secure data, cloud service providers must create a system resistant to external threats. The design and foundation of the system should also be fault-tolerant. Various factors exist in MCC, such as mobile cloud-based data exchange privacy. Until the Mobile Production Line and the Operating System establish security protocols, confidential information on mobile phones is not safe and secure. Such guidelines will aid in the protection of our system from external threats. Furthermore, the mobile device could be stolen, allowing access to highly personal information [25-29]. A security model must be established to address such challenges.

We compare all the obtainable way-outs in this category in the current study. After a comparison, the EB algorithm and the basic hashing scheme appear to be the most acceptable schemes because they are both extremely scalable. As a result, these future technologies assist us in achieving network security for mobile users. In the second section, we provide related work. The third section explains how to put the strategy into action. The specifics of the experiment are found in Section 4. Section 5 finishes with a conclusion and suggestions for future research.

II. LITERATURE SURVEY

Cloud computing is an emerging innovation that revolutionizes the way people use the Internet. In addition, cloud computing is likely to expand to mobile contexts as wireless access devices advance, with handheld sensors and devices serving as data collecting nodes. Users' anxiety about data security, on the other hand, is the critical roadblock to cloud computing are widespread adoption. These worries arise from the sensitive information stored in clouds, which are run by private cloud service providers the data owner does not trust. As a result, new highly secured topologies are required to answer customers' security concerns regarding cloud computing.

The research of Z. Zhou et al. [3] is focused on the following two directions: To protect sensing data, first use Privacy-Preserving Crypto Protocol Attribute-Based Encryption. Lightweight machines can use PP-CPABE to safely outsource heavy cryptography processes to CSPs while keeping the data private. Second, as a crypto-based access mechanism, an Attribute-Based Storage system. The performance evaluations show the solution's security and efficiency in computing, communication, and storage.

When exploring cloud computing security, Abdullah et al. [4] present an awareness using encryption and decryption approaches. If a cloud is used for data storage, encryption, and decryption, there is a possibility of unauthorized access to private information. As a result, the entire process poses a security risk. They presented a solution for analyzing how to save protected cloud data using encryption techniques. Due to the encrypted form, a hacker or unauthorized person cannot access confidential information.

Rashmi et al. [5] propose an algorithm to enhance the mobile cloud's security by ensuring data confidentiality and

integrity. To ensure stronger security, this technique employs AES and RSA. They also go over the security risks that MCC settings face. Waseem et al. [7] describe the fundamentals of MCC and the difficulties that come with it. It primarily discusses the security of data and the necessity of data security. This research looked into various data security measures so that multiple users can extensively use MCC in the future. They also suggest a mechanism to provide mobile users with access management, confidentiality, and integrity.

A scientometric study (SMS) was undertaken by Hussain et al. [9] using a set of six research topics. A total of 1711 studies were found to have been published between 2009 and 2019. After a filtering process, a selection of original studies was made. As a result, existing dangers and assaults on privacy protection were demonstrated, as well as ways to serve data. In addition, the metrics used to evaluate present privacy solutions in MCC were compiled. In addition, the domain's existing data privacy exercises was determined. Finally, there were also facets for search types and contribution types that are utilized in MCC tinted.

Arumugam et al. [10] only allow approved users to view the data in their suggested effort. Because the data was extracted unintentionally or deliberately, the attacker could not decrypt it and recover the original data. It is not possible to get the information. As a result, the RSA algorithm's methodology gives data security. When a firm decides to go to the cloud, the cloud loses its power. As a result, the level of data security offered in a straight line is proportionate to the data significance. High-performance computers and encryption are at the heart of cloud security.

Bhatia et al. [11] propose a categorization of crypto approaches and data security systems based on a novel temporal order delimitation. The poll, however, was limited to risks and attacks involving the mobile cloud. Rahimi et al.[12] studied numerous MCC security systems, most of which offload processor-intensive operations to the cloud. They identified many issues that CSPs must address to accomplish safety and confidentiality in the MCC environment.

Sookhak et al. [13] provide a remote server monitoring technique that maintains the integrity of a patient's stored data on the CSP. This system uses algebraic signatures, which allow auditors to quickly verify that the data of a smartphone network is in their hands. The author introduces a new information structure to make more prominent the DCT method and facilitate dynamic data updating operations. Users can execute data update procedures without loading the entire data set by employing DCT at the block level. Finally, the author implements and analyzes the work, which shows that DCT and logical signature require less computation power on the cloud than typical monitoring analysis techniques.

Qiu et al. [14] present a P2DS technique dubbed dynamic, secure data scheme, which is utilized to protect the data of smartphone subscribers from unauthorized users. Logical and active dispositive access algorithms were used in this strategy.

To guard against illegal entry, Odelu and Jin et al. [8] presented an encryption (CP-ABE) based ciphertext strategy. These schemes provide MCC access control and enable users to source entirely functional needs from devices to the cloud with less encoding and decoding. Additionally, phone

devices have access to data with greater flexibility. D Liu et al. [15] propose customized search results based on appropriate data such as the patient's past, geography, and interests while maintaining information protection on Mobile Cloud. All employed the k closest neighbor method, sophisticated attribute-based encryption search system, vector-space-based search, and bloom filter. The k nearest neighbor and bloom filter approaches are first utilized for relevance-based result ranking and successful multi-keyword searches. Furthermore, modern attribute-based keyword search algorithms and vector-space-based search algorithms facilitate searching with multiple users requesting massive amounts of data.

Mollah et al. [16] describe a secure method for searching cloud data via mobile devices. This system employs advanced security mechanisms such as searchable secret key encryption, digital signatures, secret key encryption, and public key encryption to safeguard data on the mobile cloud. Momeni [17] describes mobile cloud computing in detail, including definitions, architecture, reasons for development, and future research potential. Before describing mobile cloud computing, it is necessary to describe it. The authors define MCC as an infrastructure that allows data storage and processing. Integration of the mobile web with the cloud is what they call it. According to the research, on-demand solutions, low cost, resilience, and adaptability are among the benefits of MCC. They also address the issue raised by the list: security and privacy are the most pressing issues that need to be addressed.

Qijun et al. [18] looked at the latest research and developments in safe MCC. The very first examined three separate cloud infrastructures that could be used to enable future cloud-based MCC scenarios. They proved that by combining the benefits of cloud computing and mobile devices into a single system. They then looked at several issues that affect MCC infrastructures' availability and integrity. They showed that invaders could access traditional client-server systems in an MCC environment. Finally, recently proposed defense mechanisms for safeguarding MCC systems and applications were presented. The restriction simply prepares secure architecture and does not safeguard stored files or mobile devices. It has a time limit for obtaining data.

III. SYSTEM MODEL

In MCC, the suggested model employs the EB algorithm. The proposed system is a cloud-based browser application that runs on the EB algorithm based on cryptographic techniques. The cloud service customer can also use this platform to handle and transfer data to a cloud computing system. The system's process is depicted in Fig. 2. The administrator transfers a document to the cloud and encrypts the data with the EB algorithm before being placed in the dataset. Cryptographic keys are created for the decryption procedure at the very same time.

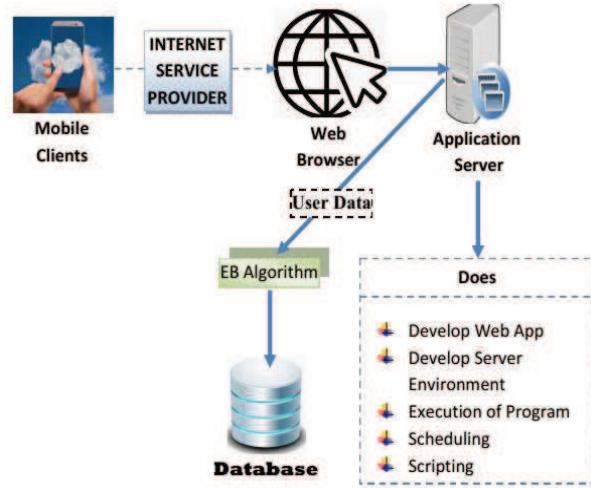


Fig. 2. Design Model

Before decrypting the ciphertext, the user needs the file's cryptographic keys and id, which are transferred back and forth between the client and the server. When a customer registers to the website with their email and password, the secret key is provided through email during the user access procedure. The client can then decode the encrypted message that uses the cryptographic keys and obtain the plaintext. If the one-time code and document id are not the same, the message will be generated incorrectly.

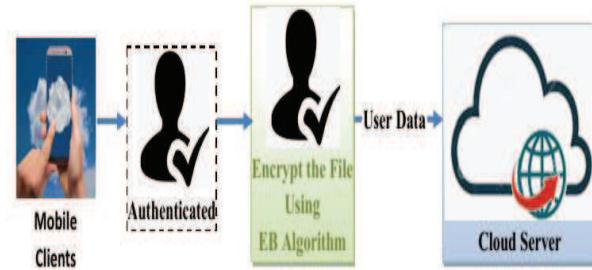


Fig. 3. Elements in the system's development

The suggested scheme shows how the process works. Clients who register to view, administer, and save their information on the online service are referred to as external users, and they can connect and interact via an internet browser. The browser communicates to server and is used in server and web environments. It can also assist with program execution, planning, and programming [30]. This will run the EB encryption algorithm, which is required to protect the user's data. When a client needs their information, the databases are typically used to store and retrieve data that are moved through the application server. The data must be extracted from the databases and done online via the application server is depicted in Fig. 3,4.

A corporation can keep massive customer and employee data records and other administration system information in both cloud environments. We've used the following process to save and extract information for any record system [31]. The administrator accepts any file from an authorized mobile client and encrypts it with the EB algorithm before storing it in the cloud. Clients access their files after decrypting them with bits of help from the server and file id.

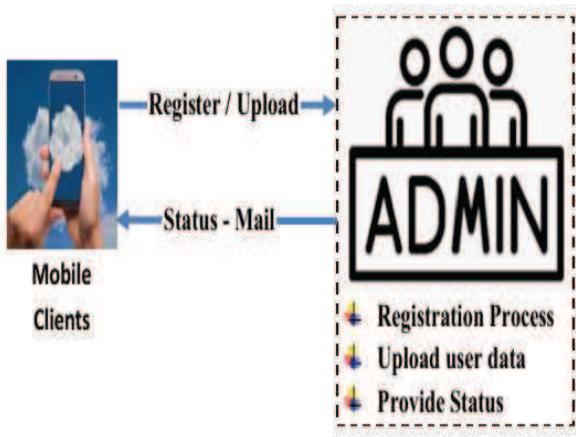


Fig. 4. Registration and Uploading Process

The EB Algorithm is a method for calculating the probability of a particular event. Our suggested method is a symmetric cryptographic algorithm, which means it encrypts and decrypts communications using the same secret key. Because it uses the same key, it requires additional security against cybercriminals and intruders [32]. Encryption of messages necessitates the use of a key (K). The cryptographic approach is implemented using an EB algorithm. The encrypted text is sent to the cloud using HTTPS, a secure connection protocol. However, the cloud service provider is aware of an incomprehensible message to an attacker. The attacker can recover the encrypted message, but the key is impossible.

With the same cloud infrastructure, these give the data owner a different type of protection for their data stored in the cloud. The block size is 64 bits, and messages over 8 bytes are ignored. It is divided into two sections: key extension and content protection.

A. Key-expansion

The input key is split into many 4168-byte subarrays. The P_a array is made up of 18 boxes (32-bit), while the S-boxes are made up of 4 arrays (32-bit) having 256 elements each one. The key's first 32 bits are XOR with P_{a1} in the next phase. The next 32 bits of a secret are XOR with P_{a2} , until every one of 448 bits are used.

B. Data encryption

The sensitive data is used with 64-bit original message and then encoded to 64-bit encrypted message at this stage. As left and right elements, it is divided into two 32-bit sections. Then, with ease, finish the XOR for identically left bits and 32 bit middle sections. This technique continues until the 16th round is completed.

C. Encryption of data and the F_n function

It transmits four S-boxes (32-bit), everyone with 256 elements apiece. The ER computation essential 32 bit of a gap halves are divided into four 8 bit blocks, p, q, r, and s. Eq. shows the modus operandi for the function F_n (1)

$$F_n = ((S_1, p + S_2, q \bmod 232) \oplus S_3, r) + S_4, s \bmod 232 \quad (1)$$

IV. EXPERIMENTAL RESULTS

This part is where performance review investigations are carried out. The proposed technique is implemented on the Java programming language. The throughput, as well as the encryption and decryption times, are evaluated. During analyzing information, every one of these solutions is not just to safeguard data. The proposed system encrypts and decrypts various input files of varied sizes using an ER algorithm (in kb). This method was developed for security purposes and to speed up the execution of "encryption and decryption procedures." The security of the ER encryption technique is sorely tested.

Metrics of performance- The performance indicators include key size, decryption time, and throughput.

A. Encryption Time

This is when converting plain text to ciphertext. The efficiency is calculated using encryption time in milliseconds. It refers to how long cryptographic algorithms take to encrypt data. The faster the encryption time, the better the algorithm's performance.

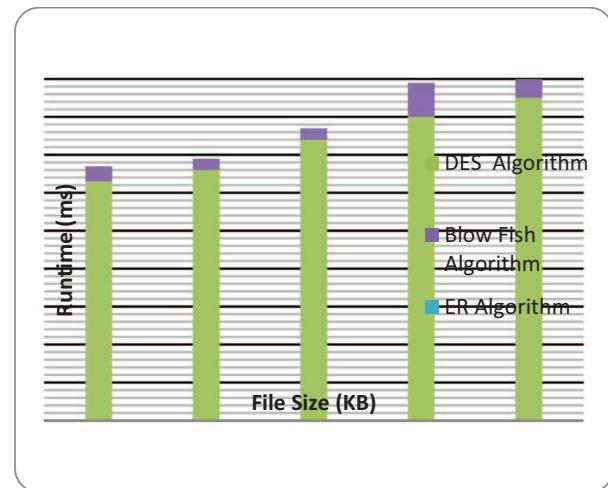


Fig. 5. Encryption Time

B. Decryption Time

The time taken for data encryption to convert encrypted messages to original messages is known as decryption time. Decoding duration is used to estimate the overall efficiency of the decoding operation. To put it another way, it shows how rapidly the decoding operation is accomplished. In most cases, decoding times are measured in milliseconds. The greater the algorithm's performance, the quicker the decryption period.

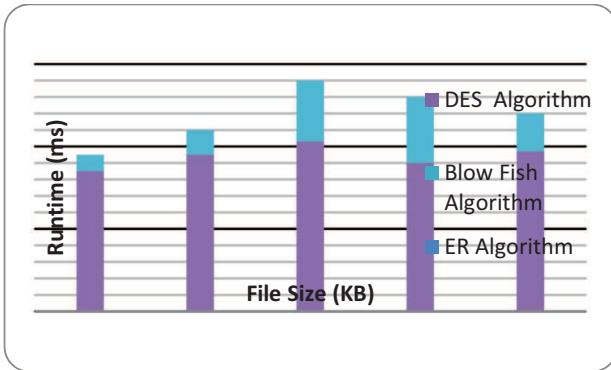


Fig. 6. Decryption Time

C. Throughput

The cryptography scheme's efficiency is measured as the proportion of entire clear text to encryption or decryption time. The greater the throughput, the better an encryption process is at protecting sensitive data.

$$\text{Throughput} = T_{pt} (\text{Kbs}) / E_{time} (\text{Ms}) \quad (2)$$

Where; T_{pt} : sum of PlainText (Kbytes)

E_{time} : Time taken for Encryption (Milliseconds).

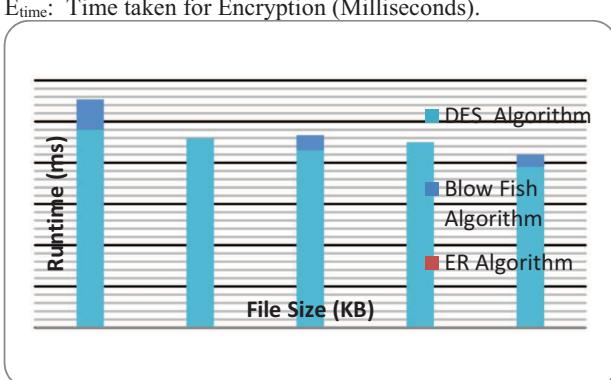


Fig. 7. Throughput

In this research, we used the ER crypto encrypting method to create a data protection model for mobile computing. Usage of mobile gadgets can save files free worry of being attacked. Because the suggested approach ensures the privacy and anonymity of any users who save their data in the cloud.

V. CONCLUSION

As a result of the preceding analysis, we can infer that mobile cloud computing has enabled phone users to conveniently protect their data on portable clouds and has improved their work easier because Users can use the cloud-based services anywhere in the world thanks to the internet. However, there are many considerations to consider as we move closer to MCC. For mobile computing, we presented a safe and privacy-preserving data storage architecture. Our technology, in particular, available — wireless devices to safely process and store information in the cloud platform at a low cost.

The suggested EB algorithm assures data and application secure operation in the mobile cloud. It also performs user authentication to authenticate user's access to a mobile cloud application. To evaluate the outcomes and assess their

security, the suggested technique might be implemented in a software program. In our recommended solution, only the authorized people get control over the data. The data cannot be decrypted unless the attacker has the encrypted cryptographic keys and file id provided by the central server. We made a fair evaluation of ER, DES, and Blowfish. The simulations revealed that the ER method outperformed DES and the blowfish algorithm in the encryption, decryption, and throughput.

REFERENCES

- [1] Kandavel, N., and A. Kumaravel. "A Novel Royal Seal Cloudlet for Security Enhancement in Mobile Cloud Computing." International Journal of Computer Science and Information Security (IJCSIS) 17, no. 1 (2019).
- [2] Bhavya Shree M N, K S Ananda Kumar, Kavyashree S, Gagana H, S Geetha; Accessing Data In Cloud Platform Using Identity And Providing Strong Security, Auditing Based On IBDO Scheme; Global Journal of Engineering Science and Researches [ICRTCET-2018]; 2018, pp 54-61.
- [3] Z. Zhou and D. Huang, "Efficient and secure data storage operations for mobile cloud computing," 2012 8th international conference on network and service management (cnsm) and 2012 workshop on systems virtualization management (svm), 2012, pp. 37-45.
- [4] Abdullah et al., "The Secure Data Storage in Mobile Cloud Computing," Computer Engineering and Intelligent Systems, Vol.6, No.5, 2015.
- [5] Rashmi et al., "Data Security in Mobile Cloud Computing," International Conference on Sustainable Computing in Science, Technology & Management, (SUSCOM-2019), 2019.
- [6] J. A. Shanny and K. Sudharson, "User preferred data enquiry system using mobile communications," International Conference on Information Communication and Embedded Systems (ICICES2014), 2014, pp. 1-5, doi: 10.1109/ICICES.2014.7033943
- [7] Waseem, M., Lakhani, A. and Jamali, I. (2016) Data Security of Mobile Cloud Computing on Cloud Server. Open Access Library Journal, 3, 1-11. doi: 10.4236/oalib.1102377.
- [8] K. Sudharson, Ahmed Mudassar Ali, A.M. Sermakani, "An organizational perspective of knowledge communication in developing entrepreneurship education for engineering students," Procedia - Social and Behavioral Sciences, vol. 73, pp. 590-597, 2013, doi:https://doi.org/10.1016/j.sbspro.2013.02.095.
- [9] Hussain et al., "Privacy and data protection in mobile cloud computing: A systematic mapping study," PLOS ONE, 15(6):e0234312. https://doi.org/10.1371/journal.pone.0234312 June 11, 2020.
- [10] Arumugam et al., "Secure data sharing for mobile cloud computing using RSA," IOP Conf. Series: Materials Science and Engineering 1055 (2021).
- [11] Bhatia T.,& Verma A. K.(2017).Data security in mobile cloud computing paradigm: a survey, taxonomy and openresearchissues.TheJournalofSupercomputing,73(6),2558–2631.
- [12] RahimiM.R.,RenJ.,LiuC.H.,VasilakosA.V.,&VenkatasubramanianN. (2014). Mobile cloud computing:ASurvey, state of art and futuredirections.MobileNetworksandApplications, 19(2),133–143.
- [13] Chandni Patel, S. C. and Patel, B. "A Data Security Framework for Mobile Cloud", International Journal of Advanced Research in Computer and Communication Engineering, pp. 254-257, 2015.
- [14] Ibtihal Mouhib, El Ouardighi Driss and Khalid Zine-Dine "Encryption as a service for securing data in mobile cloud computing", 15th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 546-55, 2015.
- [15] Hongwei Li, D. L, and Yuanshun Dai, S. Y. "Personalized Search Over Encrypted Data With Efficient and Secure Updates in Mobile Clouds", IEEE Xplore, pp. 1-12, 2015.
- [16] Muhammad Baqer Mollah and M. A. "Secure Data Sharing and Searching at the Edge of Cloud-Assisted Internet of Things". IEEE Cloud Computing, pp. 34-42, 2017.
- [17] M. R. Momeni et al., "A Survey of Mobile Cloud Computing: Advantages , Challenges and Approaches," vol. 15, no. 4, pp. 14–28.

- [18] Gu, Qijun, and Mina Guirguis. "Secure mobile cloud computing and security issues." In *High Performance Cloud Auditing and Applications*, pp. 65-90. Springer, New York, NY, 2014.
- [19] Li, R., Shen, C., He, H., Gu, X., Xu, Z. and Xu, C.Z., 2017,A lightweight secure data sharing scheme for mobile cloud computing,IEEE Transactions on Cloud Computing, 6(2), pp.344-357.
- [20] Boomathi, V. and Banumathi, A., 2020, Efficient analysis of encryption algorithms RSA, DES and AES for sensitive data, Purakala with ISSN 0971-2143 is an UGC CARE Journal, 31(4), pp.1937-1945.
- [21] J. Aruna Jasmine, V. Nisha Jenipher, J. S. Richard Jimreeves, K. Ravindran and D. Dhinakaran, "A traceability set up using Digitalization of Data and Accessibility." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020, pp. 907-910, doi: 10.1109/ICISS49785.2020.9315938.
- [22] S. Arun and K. Sudharson. "DEFECT: discover and eradicate fool around node in emergency network using combinatorial techniques." *Journal of Ambient Intelligence and Humanized Computing*, 1-12, 2020. doi: <https://doi.org/10.1007/s12652-020-02606-7>.
- [23] Sinha, D., Datta, S. and Das, A.K., 2020,Secure Data Sharing for Cloud-Based Services in Hierarchical Multi-group Scenario,in Advances in Computational Intelligence, pp. 229-244, Springer, Singapore.
- [24] K. Sudharson and V. Parthipan, "A Survey on ATTACK – Anti terrorism technique for adhoc using clustering and knowledge extraction," *Advances in Computer Science and Information Technology. Computer Science and Engineering. CCSIT 2012. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, Berlin, Heidelberg, pp 508-514, vol 85, 2012. doi:https://doi.org/10.1007/978-3-642-27308-7_54.
- [25] Hu, Y., Kumar, S. and Popa, R.A., 2020, Ghostor: Toward a secure data-sharing system from decentralized trust,in 17th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 20), pp. 851-877.
- [26] N.Partheeban, K.Sudharson and P.J.Sathish Kumar, "SPEC- Serial property based encryption for cloud", *International Journal of Pharmacy & Technology*, Vol. 8, No. 4, pp. 23702-23710, 2016.
- [27] Alnajrani, H.M., Norman, A.A. and Ahmed, B.H., 2020, Privacy and data protection in mobile cloud computing: A systematic mapping study, *Plos one*, 15(6), p.e0234312
- [28] Ibtihal, Mouhib, and Hassan Naanani, 2020,Homomorphic encryption as a service for outsourced images in a mobile cloud computing environment, In *Cryptography: Breakthroughs in Science andPractice*, pp. 316-330. Global IGI.
- [29] K.Sudharson, Ahmed Mudassar Ali and N.Partheeban, "NUITECH – Natural user interface technique foremulating computer hardware", *International Journal of Pharmacy & Technology*, Vol. 8, No. 4, pp. 23598-23606, 2016.
- [30] Dhinakaran D., Joe Prathap P.M. (2022) Ensuring Privacy of Data and Mined Results of Data Possessor in Collaborative ARM. In: Ranganathan G., Bestak R., Palanisamy R., Rocha Á. (eds) *Pervasive Computing and Social Networking. Lecture Notes in Networks and Systems*, vol 317. Springer, Singapore. https://doi.org/10.1007/978-981-16-5640-8_34.
- [31] Kumar, Y.K. and Shafi, R.M., 2020,An efficient and secure data storage in cloud computing using modified RSA public key cryptosystem, *International Journal of Electrical and Computer Engineering*, 10(1), p.530.
- [32] Hidayat, T. and Mahardiko, R., 2020. A Systematic Literature Review Method On AES Algorithm for Data Sharing Encryption On Cloud Computing, *International Journal of Artificial Intelligence Research*, 4(1).

Hybrid filter and wrapper methods based feature selection for crop recommendation

Madhuri J
Department of Computer science and Engineering
Bangalore Institute of Technology ,Bengaluru
madhurij@bit-bangalore.edu.in

Dr. Indiramma M
Department of Computer science and Engineering
B.M.S. College of Engineering, Bengaluru
indira.cse@bmscse.ac.in

Abstract

The agriculture sector is certainly benefited from advances in science and technical proficiency, which has resulted in the generation of vast volume of data. With huge processing techniques, machine learning has enabled perception of new possibilities in agriculture management. Implementing machine learning algorithms over agricultural data is the significant onset for investigating and determining solutions to agrarian challenges. Agricultural data emerges from diverse sources such as weather, soil, crop characters and so on. Feature selection methods plays significant role to eliminate irrelevant features and identify significant features thereby enhancing machine learning model's performance. The paper proposes a hybrid feature selection strategy by combining filter and wrapper methods. The proposed method obtains optimal features from soil properties, crop characteristics, and climatic parameters in order build a crop recommendation model with better accuracy and performance. The effectiveness of the model is measured by considering all the features in the dataset and with the features obtained from proposed method. The proposed feature selection method is validated through evaluation metrics MSE, RMSE, MAE, R². The machine learning models viz. artificial neural networks and decision tree are implemented with the selected features.

Keywords: *Crop recommendation, feature selection, agriculture, soil, climate, ArNN*

I. INTRODUCTION

The innovations in the field of science and technology has generated massive amounts of data in the agrarian sector which can be effectively made use for recommendation of suitable crop, crop yield prediction, fertilizer recommendation, plant disease classification, and so on[1]. Machine learning is the effective field in analyzing the data related to farm, crop and climate for the recommendation of suitable crop for the land. The performance of machine learning model developed for crop recommendation has to be enhanced for accomplishment of superior results. Existing ways to determine crop recommendation based on land suitability are time consuming and expensive due to the high number of inputs required[2]. As a result, it's critical to create methods for minimizing and optimizing the input parameters

for crop recommendation. Machine learning model's performance can be boosted by optimization of potential input parameters by feature selection[3]. The advantages of feature selection are as follows: (a) reduction in overfitting: as the redundant features and noise is removed, the machine learning model would perform better with test data. (b) Reduction in training time: as the dimensionality of the data is reduced, the recommendation model will be computationally fast. (c) Increase in model accuracy: as the irrelevant features are removed, accuracy of the model will be increased.[4]

Precision agriculture is the term deliberated for applying location specific agriculture method and it is the amalgamation of farm management and information technology. Integrating machine learning with precision agriculture help build a crop management system that supports farming activities such as choosing suitable crop, fertilizer management, irrigation management and crop yield prediction. As division of efficient crop management system, a recommendation of the suitable crop considering climatic parameters, soil and crop characteristics is necessary to benefit the farmers [2]. The recommendation of the crops are based on suitability measures 1.High. 2. Moderate 3. Marginal 4. Not suitable[5]. But, when all the features are considered for the model, space complexity and the computational time would be increased. Feature selection is significant as the unnecessary features can influence the ML model's performance negatively. The aim of this paper is to develop a hybrid feature selection strategy by combining filter methods and wrapper methods to identify suitable features from climate, soil and crop characteristics conductive to build efficient crop recommendation model.

II. LITERATURE REVIEW

Feature selection is the method for identifying suitable input parameters using various statistical and data mining techniques to enhance the model's performance. The goal of feature selection to delete non-essential features without impacting learning performance. The methods used for feature selection methods have number of advantages, including lowering data acquisition costs and making classification models more understandable[6]. The data subset obtained should be evaluated using different metrics to discriminate and obtain optimal classifier. [7] Specifies different evaluation

metrics to be considered while building the classifier with feature subset. Among the different feature selection methods used the important are (1) Filter methods: uses ranking technique to order the variables for selection (2) Wrapper methods: evaluates every subset of data and chooses the subset with high performance. (3) Embedded methods [8]. Wrapper approaches test the relevance of a subset of features by actually training a model on it, whereas filter methods measure the importance of features by their correlation with the dependent variable. The filter methods use a ranking system to determine the significance of features and considers to remove low scoring features. The filter approaches are identified to be quick, scalable, computationally simple, and classifier-independent but they don't consider the influence of selected features on the efficiency of the model[9]. The wrapper methods work in a similar way to the filter methods, but instead of using an independent measure for subset evaluation, they employ a predefined classification algorithm. The accuracy of the model is measured by using the subset of the features and are evaluated. When compared to filter methods, wrapper approaches produce superior results, but they are more computationally expensive when the number of available features is very large[9] [10]. Hence we propose a hybrid method which utilizes the advantages of both filters and wrappers feature selection. Feature selection methods are used in the field of precision agriculture to increase the efficiency of the model. Modified genetic algorithm was developed to choose preeminent features that would enhance the performance of regression, artificial neural networks and adaptive neuro fuzzy inference system algorithms. The performance of crop yield prediction model increased with feature selection and extraction strategies[11]. [12]evaluates the performance of different feature subsets using different feature selection methods for crop yield prediction.

III. PROPOSED WORK

This section presents the entire methodology, involving feature selection strategy and the ML algorithm formulated to provide crop recommendation. The figure 1 shows the hybrid feature selection method based crop recommendation system.

The methodology consists of 3 stages. In stage 1 data preprocessing is carried out to predict the missing values and eliminate the outliers. The missing values are replaced with the average value of the data. In the stage 2, hybrid feature selection method was employed to select the most relevant values. The hybrid feature selection method combines correlation based filter method and backward elimination wrapper method. A filter method can provide reduced feature subset to wrapper method, which uses prediction accuracy to evaluate the feature subset. The hybrid feature selection method consists of two phases: Firstly, the filter method reduces the number of features in the dataset by eliminating unimportant features. Secondly, the wrapper method finds the ideal feature subset from reduced dataset. In stage 3, the selected features are used to build the machine learning model that recommends the location specific crop.

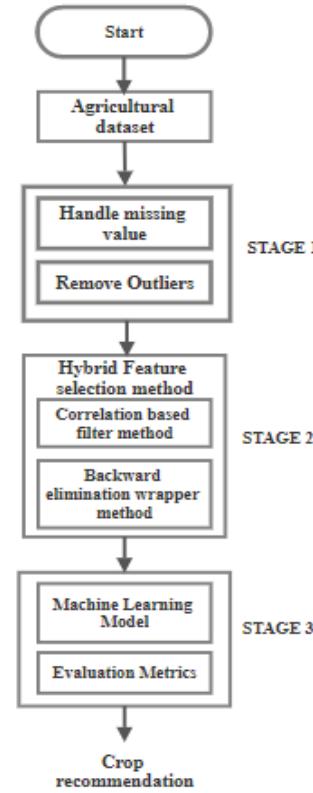


Figure 1: Proposed hybrid feature selection methodology

A. Dataset description

In this work, we focus on obtaining datasets for the Doddaballapur (dist.) in Karnataka, India for the locations Hadonahalli and Durgenahalli. The district is located at latitude $13^{\circ} 20'$ north and longitude $77^{\circ} 31'$ east. The data set includes unique soil and climate features, along with crop characteristics of finger millet, maize, rice and sugarcane. The meteorological data is collected for the period of 2007 to 2017, from Agro meteorology Section, University of Agricultural Sciences, Bengaluru. The land/ soil characteristics for this location is obtained from National Bureau of Soil Survey and Soil Usage Planning (NBSS & LUP), Bengaluru.

	Feature code	Feature	Feature description
1	St	Soil texture	
2	SpH	Soil pH	
3	Sgc	Gravel code	
4	Sec	Erosion code	
5	Ssl	Slope	

6	Sdr	Soil drainage	
7	Sdp	Depth	
8	Spt	Soil potassium	Potassium present in the soil
9	Sni	Soil nitrogen	Nitrogen present in the soil
10	Sph	Soil phosphorus	Phosphorus present in the soil
11	LGP	Length of growing period	Time taken from sprouting to harvest of the crop
12	Erd	Effective root depth	Rooting distance of the crop from surface
13	Tmin	Minimum temperature	Average minimum temperature
14	Tmax	Maximum temperature,	Average maximum temperature
15	Hum	Humidity	Amount of water vapour in the air
16	WS	Wind speed	Rate of wind blow
17	SSH	Sunshine hours	Measure of daytime duration

Table 1: Dataset attributes and its description.

Crop suitability depends on various features that is categorized as soil properties, crop characteristics, and climatic parameters. The data of soil, climate and crop characteristics were collected for the study. Soil properties include soil texture, soil pH, gravel code, erosion code, slope, soil drainage, depth, soil potassium, soil nitrogen, soil phosphorus. Crop characteristics include length of growing period, effective root depth and mean temperature. Climatic parameters include minimum temperature, maximum temperature, humidity, wind speed, sunshine hours, potential Evapotranspiration. The crop to be recommended is the dependent variable whose value has to be predicted[13].

B. Stage 1: Data preprocessing

Handling missing values: Incomplete data causes inaccuracy in ML models predictions. The missing values are replaced with the mean or average value of that feature[14][15].

Outlier elimination: The extreme values in the dataset are removed as the variability in the data may decrease accuracy in the predictions.

C. Stage 2: Feature selection

C1. Filter Method: Correlation-Based Feature Selection method

Correlation based feature selection is a simple filter method that ranks the features using a heuristic evaluation function based on correlation. Pearson correlation coefficient is used for the process. It is based on the concept that a good attribute has a strong correlation with the classes and does not correlate with other attributes. The attribute's strong correlation with other attributes implies that it is redundant. An irrelevant attribute is defined as one that has a low correlation to the class. Henceforth the duplicate and irrelevant attributes are removed[16]. Only the features with smallest correlation values with other features are selected and added into the feature list. The Pearson correlation coefficient is given by,

$$r = \frac{K r_{ci}}{\sqrt{K+K(K+1)r_{ii}}} \quad (1)$$

Where K represents the number of features in the dataset, r_{ci} is the correlation between each feature and the class c, r_{ii} is the correlation between the features.

C2. Wrapper Method: Backward-elimination feature selection (BEFS)

The backward elimination method begins with dataset obtained after employing filter method. The wrapper method removes the feature whose removal improves classification accuracy the most. This operation is performed until eliminating any of the remaining features has no effect on classification performance[17]. BEFS consists of four major steps: Step 1: Selecting the significant level for p-value. Usually p-value is chosen as 0.05.

Step 2: Fit the ML model with all the features and check the accuracy of the model.

Step 3: Find the feature with high P-value, if the P-value is greater than the significant value set in step 1, then the feature is removed from the dataset.

Step 4: If any feature is removed then repeat step 2 and 3. Repeat the steps till the p-value of all the features remain below the significant value chosen.

D. Stage 3: Machine Learning Algorithms

The proposed hybrid feature selection method is implemented with the following ML algorithms:

- Artificial neural networks
- Decision tree

Artificial neural networks (ArNNs) are complex mathematical learning models inspired by biological neural networks. The ArNN consists of three layers namely, input layer, hidden layer and output layer. It can consist of multiple hidden layers. Neurons are the basic building blocks of ArNN and the

activation function attached to every neuron decides the output to be produced[18]. The ArNN is trained with Backpropagation algorithm in the present research. ReLU activation function is used for the hidden layers. As the output of the recommendation model is multiclass classifier, softmax activation function is used in output layer[19]. The ID3 approach is used to train the decision tree classifier on the same dataset. The recommendation model in several published papers has been based on a decision tree classifier. Hence, a decision tree is employed to compare the findings of the ArNN.

Evaluation Metrics

Evaluation metrics are used to measure model performance[7]. The metrics are used in this study to evaluate the efficiency of hybrid feature selection model and after. The various metrics considered for the evaluation in this study are:

- Mean square error (MSE): This defines the difference between the prediction of the model and the actual output. The difference is squared and averaged across the entire dataset. The formula for MSE is as shown in equation 2, where P is the predicted value and A is the actual value.

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_i - A_i)^2 \quad (2)$$

- Mean Absolute error (MAE): This defines the difference between the prediction of the model and the actual output. The absolute value of the difference is taken and squared across the entire dataset.

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - A_i| \quad (3)$$

- Root mean square error (RMSE): It measures the standard deviation of the prediction errors. The formula of RMSE is given by equation 4.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - A_i)^2}{N}} \quad (4)$$

- R² or correlation coefficient: It is the measure used to determine the correlation between predicted and actual values. The formula of RMSE is given by equation 5.

$$R^2 = \frac{n(\sum PA) - (\sum P)(\sum A)}{\sqrt{n[\sum P^2 - (\sum P)[\sum A^2 - (\sum A)]}}} \quad (5)$$

IV. RESULTS AND DISCUSSION

His section describes the results of the hybrid feature selection method is evaluated with ML algorithms described in section 3.3. The most relevant or significant features are selected from the dataset using feature selection in order to enhance the model's performance. Python was employed as the coding language on Windows 10 environment. The experiment was

conducted on the combination of soil, climate and crop datasets. The features of the dataset are described in Table 1.

The correlation based feature selection method creates the correlation score for each independent feature and dependent class. The subset with highest correlation score was designated as the final set. Totally 15 features are selected after this filter stage viz. St, SpH, Sgc, Sec, Sdr, Sdp, Spt, Sni, Sph, LGP, Tmin, Tmax, Hum, WS, SSH.

The backward elimination feature selection is based on the p-value. The p-value is determined for all the features obtained from filter method against the dependent variable. The features selected after this phase are St, SpH, Sgc, Sec, Sdr, Sdp, Spt, Sni, Sph, LGP, Tmax, Hum, WS, SSH. The present work implements the hybrid feature selection method over the Artificial Neural Networks (ArNN) and Decision tree algorithms to build the crop recommendation model. The ArNN was built with 3 hidden layers and output layer consists of 4 neurons representing the suitability classes for rice, maize, finger millet and sugarcane. ArNN is trained using Backpropagation algorithm for 89 epochs[13]. Decision tree is built using ID3 algorithm. Suitability is recommended in terms of class: High suitability-Class 1, Moderate suitability-Class 2, Marginal suitability- Class 3, not suitable- Class 4. The efficiency of the models is measured with:

(1) The models are built with all the features in the dataset and crop recommendation results are predicted.

(2) The models are built with the features selected by utilizing the proposed hybrid feature selection method. The recommendation results are evaluated using the statistical measures mentioned in the previous section.

. The developed model subjected to evaluation metrics mentioned in section 3.4. The values of the metrics are as shown in Table 2 and Table 3.

The results of the evaluation measures calculated for the machine learning models with all the features in the dataset and with features selected from hybrid method are shown graphically in figure 2a and 2b.

Algorithm	Performance metrics considering all the features of dataset			
	MSE	MAE	RMSE	R ²
ArNN	0.092	0.102	0.186	0.42
Decision tree	0.298	0.390	0.432	0.31

Table 2: Performance evaluation of ML models considering all the features in the dataset

Algorithm	Performance metrics considering the features obtained from hybrid feature section method			
	MSE	MAE	RMSE	R ²
ArNN	0.071	0.182	0.162	0.66
Decision tree	0.228	0.176	0.322	0.38

Table 3: Performance evaluation of ML models considering the features obtained from hybrid feature section method

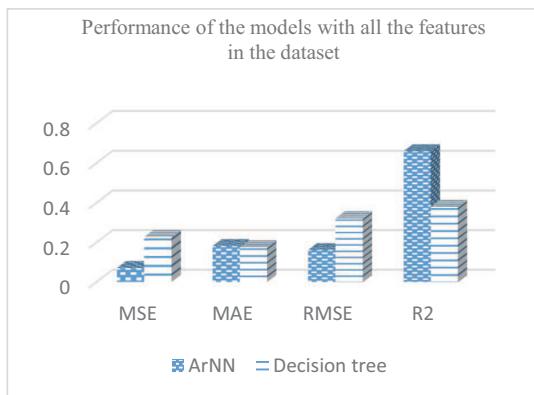


Fig 2a: Performance of machine learning models considering all the features

The evaluation metrics define the performance of the developed model. It is evident from the results obtained that the machine learning models are performing better with hybrid feature selection method.

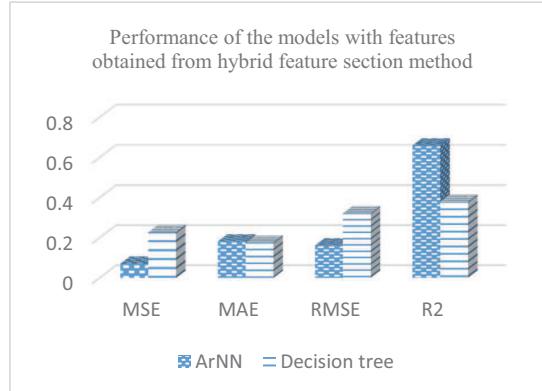


Fig 2b: Performance of machine learning models considering the features from hybrid feature selection method

V. CONCLUSION

The proposed filter and wrapper based hybrid feature selection method selects optimal features to build an efficient crop recommendation model. The hybrid method combines correlation based feature selection and backward elimination wrapper method for feature selection. Experiments was conducted on the soil, climate and crop datasets. The important features selected are well-defined and positively affects the recommendation model.

Experimental results obtained confirm that performance of crop recommendation model is better with the features selected hybrid feature selection method. The efficient crop recommendation model developed would benefit the farmers in choosing the right crop depending on their location. Further the proposed method can be tested over other agricultural datasets such as pricing data, fertilizer data and irrigation decision models. Features remote sensed image data can also be considered for developing agricultural recommendation systems.

REFERENCES

- [1] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Comput. Electron. Agric.*, vol. 151, no. May, pp. 61–69, 2018, doi: 10.1016/j.compag.2018.05.012.
- [2] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," *2016 8th Int. Conf. Adv. Comput. ICoAC 2016*, pp. 32–36, Jun. 2017, doi: 10.1109/ICOAC.2017.7951740.
- [3] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/J.NEUROCOM.2017.11.077.
- [4] L. Herrera and A. I. Negueruela, "Mutual Information-based Feature Selection in Spectrometric Data for Agriculture Applications Machine Learning applied to Cosmic Rays View project Multidisciplinary Applications View project," Accessed: Feb. 02, 2022. [Online]. Available: <https://www.researchgate.net/publication/282987273>.
- [5] A. AL-Taani, Y. Al-husban, and I. Farhan, "Land suitability evaluation for agricultural use using GIS and remote sensing techniques: The case study of Ma'an Governorate, Jordan," *Egypt. J. Remote Sens. Sp. Sci.*,

- [6] vol. 24, no. 1, pp. 109–117, Feb. 2021, doi: 10.1016/J.EJRS.2020.01.001.
- [6] E. Cantú-Paz, “Feature Subset Selection, Class Separability, and Genetic Algorithms,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3102, pp. 959–970, 2004, doi: 10.1007/978-3-540-24854-5_96.
- [7] M. Hossin and Sulaiman, “A REVIEW ON EVALUATION METRICS FOR DATA CLASSIFICATION EVALUATIONS,” *IJDCKP) Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, 2020, doi: 10.5121/ijdckp.2015.5201.
- [8] G. Chandrashekhar and F. Sahin, “A survey on feature selection methods,” *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014, doi: 10.1016/J.COMPELECENG.2013.11.024.
- [9] S. A. Rahman and S. Fong, “Feature selection methods: Case of filter and wrapper approaches for maximising classification accuracy,” *Artic. Pertanika J. Sci. Technol.*, 2018, Accessed: Feb. 06, 2022. [Online]. Available: <http://www.pertanika.upm.edu.my/>.
- [10] Y. Piao, M. Piao, K. Park, and K. H. Ryu, “An ensemble correlation-based gene selection algorithm for cancer classification with gene expression data,” *Bioinformatics*, vol. 28, no. 24, pp. 3306–3315, Dec. 2012, doi: 10.1093/BIOINFORMATICS/BTS602.
- [11] K. Aditya Shastry and H. A. Sanjay, “A modified genetic algorithm and weighted principal component analysis based feature selection and extraction strategy in agriculture,” *Knowledge-Based Syst.*, vol. 232, p. 107460, Nov. 2021, doi: 10.1016/J.KNOSYS.2021.107460.
- [12] P. S. Maya Gopal and R. Bhargavi, “Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms,” *Appl. Artif. Intell.*, vol. 33, no. 7, pp. 621–642, 2019, doi: 10.1080/08839514.2019.1592343.
- [13] J. Madhuri and M. Indiramma, “Artificial Neural Networks Based Integrated Crop Recommendation System Using Soil and Climatic Parameters,” *Indian J. Sci. Technol.*, vol. 14, no. 19, pp. 1587–1597, 2021, doi: 10.17485/ijst/v14i19.64.
- [14] V. A. Zamfir, M. Carabas, C. Carabas, and N. Tapus, “Systems monitoring and big data analysis using the elasticsearch system,” *Proc. - 2019 22nd Int. Conf. Control Syst. Comput. Sci. CSCS 2019*, pp. 188–193, May 2019, doi: 10.1109/CSCS.2019.00039.
- [15] N. Mittag, “Implications: Benefits, Risks and a Method for Missing Data,” 2013.
- [16] N. J. Cornish *et al.*, “Preprocessing Using Correlation Based Features Selection on Naive Bayes Classification The effects of orbital motion on LISA time delay interferometry Modeling of Parity Status of The Mother and Basic Immunization Giving to Infants with Semiparametric Bivariate Probit (Case Study: North Kalimantan Province in 2017) Preprocessing Using Correlation Based Features Selection on Naive Bayes Classification,” doi: 10.1088/1757-899X/982/1/012012.
- [17] “No Title,” [Online]. Available: Srinidhi, S. (2019). Backward Elimination for Feature Selection in Machine Learning. Towar. Data Sci.
- [18] A. Mehnatkesh, S. Ayoubi, A. Jalalian, and A. Dehghani, “Prediction of rainfed wheat Grain yield and biomass using Artificial Neural Networks and Multiple Linear Regressions and determination the most factors by sensitivity analysis,” no. June 2015, 2009.
- [19] J. Feng and S. Lu, “Performance Analysis of Various Activation Functions in Artificial Neural Networks,” *J. Phys. Conf. Ser.*, vol. 1237, no. 2, pp. 111–122, 2019, doi: 10.1088/1742-6596/1237/2/022030.

PORATBLE TIME GOVERNING SWITCH BOX USING INTERNET OF THINGS

Mr.S.Chandrasekar
Electrical and Electronics Engineering
Kongu Engineering College
Erode,Tamil Nadu
chasekar@kongu.ac.in

A.Manikandan
Electrical and Electronics Engineering
Kongu Engineering College
Erode,Tamil Nadu
manikandansmart10@gmail.com

P.Dinesh
Electrical and Electronics Engineering
Kongu Engineering College
Erode,Tamil Nadu
dineshpalanis471@gmail.com

S.Logesh
Electrical and Electronics Engineering
Kongu Engineering College
Erode,Tamil Nadu
logeshsubramani2001@gmail.com

B.Hari Prasath
Electrical and Electronics Engineering
Kongu Engineering College
Erode,Tamil Nadu
bvhariprasath2001@gmail.com

Abstract-In today's world, with the help of rapid technological advancement, we can easily control the home appliances at our fingertips with the help of the Internet of Things (IoT). It plays a vital role in emerging technological development. Generally, most people will forget to turn off their household appliances regularly. For example, devices like a mobile charger, laptop charger, radio, table fan, computer, television, etc. This will result in unnecessary energy consumption and a decline in device efficiency. One of the main issues is that the standard IoT devices available in the market are bulky and expensive. They are not easily afforded by normal people. This article proposes a portable IOT home automation device to control the on-time, off-time, and working duration of the devices that can be controlled with the help of the internet. A web connection is established between the microcontroller & web server and the real-time database. Google Firebase is used as the database to store and retrieve data. The connected devices can be controlled from any place through mobile phones with the help of the internet. The major value proportions of our project are portability, user-friendly and easy affordability.

Keywords – Portable IOT devices, IOT switch box, Home Automation, Google firebase, IOT mobile application

I. INTRODUCTION

In the present world, the entire population is dependent on the operation of several industrial machines and household appliances for their comfort of living. The need for automation devices is mandatory to achieve perfect control, easy access, and high efficiency. The major two classifications of automation are industrial automation and home automation. The automation can be obtained by internet-enabled automation devices, machine learning interfaces, artificial intelligence, etc. Out of these, IoT is one of the best ways to obtain automation control of devices at low cost with high efficiency. Technological transformation is happening rapidly in the field of IoT and this results in convenience and comfort living through smart devices [3,4]. Industrial automation can be achieved by different techniques such as fixed automation, programmable automation, flexible automation, and integrated automation. When compared with industrial automation, home automation is the essential need of every human in their daily lifestyle. Household appliances can be broadly classified as pre-installed devices like fans, light bulbs, and plug-in devices like television, computer, etc.

Currently, many IoT home automation devices are available in the market. Even though they are efficient, most of them are complex and are not easily portable. The automation part is built along with the device. Eg: Smart heater, Smart television, Smart washing machine, etc. These devices are expensive and are not easily afforded by normal persons. Normal middle-class persons don't need these devices in their homes. Their only need is to automate their daily household appliances in the cheap and best way. A smart device with a feature of controlling the devices connected to it will be a better solution to this problem. So, this paper proposes a design of cost-effective, user-friendly, and easily portable home automation device for small and large plug-in devices [7]. The significant contributions of this paper are as follows,

- Low-cost and user-friendly portable home automation device is designed
- Multiple plug-in devices can be controlled easily
- Various time-based control features are provided like on/off control, run time control, and scheduled operation

The proposed design of this article will resemble a conventional switch box that can be carried to any place easily [1,2]. The controller circuit is placed inside the device and the switch ports are provided for the connection of plug-in devices. The microcontroller is the primary controller unit. The microcontroller will take less time for the operation. It is easy to use and performs multiple tasks. The maintenance is also simple. The microcontroller gets connected with the web server through the internet. The web server is one of the best approaches to obtain the necessary services by the end-users over the internet [10]. The devices which are connected to the switch port can be controlled via the internet. A mobile app is developed by using MIT app inventor and used to interact with the database. Based on the user's input, the data will be stored in the database and the operations will be performed accordingly. This process of obtaining the necessary response for the desired input is called hypertext transfer protocol (HTTP). This proposed design helps to control the connected devices from any place through mobile phones with the help of the internet. This will help normal persons to get comforted with the help of smart technologies [5].

II. EXISTING SWITCH CONTROL AND RECENT TRENDS

Internet of things is a developing field that helps users to control or monitor appliances via the internet. The switchbox-based IOT Home automation device was proposed by A.Karthi and G.Sabareesh in their paper entitled Design of Internet Controlled Switch Box using IOT [1]. They have proposed a design consisting of a prebuilt switchbox connected to the microcontroller. They have controlled the devices which are connected to the switch ports with the help of the Internet. Even though they have provided the basic on/off feature the hardware setup is not properly designed for real-time usage.

Siddharth Karanchery and N. Rakesh have also proposed a similar design in their article entitled Smart Power Socket using the Internet of Things [2]. They have connected the microcontroller with the Blynk app and provided the control features. A limitation of their work is that only on/off features are provided on the software side. We have considered these papers for our reference and developed the design to overcome the conventional difficulties.

III. PROBLEMS IN THE EXISTING METHODS

Nowadays, Automation has become an integral aspect of the world. Currently, there are numerous home automation gadgets available on the market right now. However, they are costly and out of reach for most middle-class people. Furthermore, the automation part is inbuilt along with the device. So, they are not easily carried away from one place to another. These were some of the challenges existing in the current home automation devices.

IV. PROPOSED METHODOLOGY

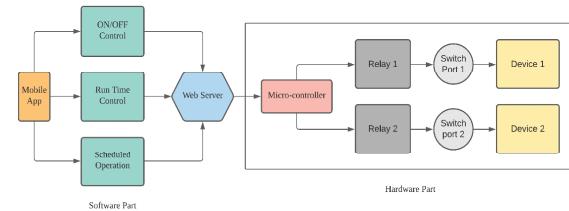


Fig. 1. Control and Power flow diagram of the portable switch box

The major objective of our design is to bring a portable home automation device to the doorsteps of every middle-class person. A low-cost, user-friendly home automation device was designed [7]. It is a versatile home automation device that works with both small and large plug-in devices like laptops, phone chargers, LED TV, radio, table fan, etc. This design consists of a microcontroller connected to a web server with the help of the internet. The mobile application was developed by using the MIT app inventor. The real-time database, Google

firebase acts as the database for the mobile application. Based on the user's input in the mobile app, the devices which are connected to the switch ports will be controlled accordingly [15]. Our design consists of three modes of operation.

- ON Time / OFF Time control
- Run Time control
- Scheduled Operation

In the on-time / off-time control, the user can turn on and turn off the device easily. The run-time control is used to set the working duration of the device and the scheduled operation helps the user to set the on-time and off-time of the device.

V. WORKING

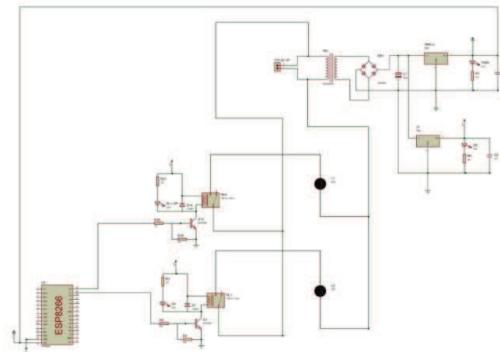


Fig. 2. Circuit Diagram of Hardware setup

The proposed design will resemble a conventional switch box. It is easily portable and occupies less space. The devices which are connected to the switch ports can be controlled from anywhere with the help of the Internet [8.9]. Initially, one part of the 230 alternating voltage supply is stepped down to 6V with the help of the transformer. It is then fed to the rectifier circuit. The rectifier circuit converts the alternating current (AC) input to direct current (DC). Then the voltage gets regulated to 5V with the help of the voltage regulator LM7805. Then it is fed to the microcontroller. The two output pins of the microcontroller are connected to the relays. The ends of the relay are connected to the switch port. The other end of the supply is also connected to the switch ports [5,6].

The Node MCU microcontroller gets connected with the web server (Google firebase) through the Internet. As soon as the user presses the necessary option in the mobile app, the data changes will get changed in the Google Firebase. These changes will be retrieved by the Node MCU and the corresponding pulse signals will be provided accordingly [11]. To obtain mobile-based control, we are using the web server and mobile app. The mobile app was developed by MIT app inventor. It is connected to the web

server with the help of the internet. Once the client raises the request (mode selection) the corresponding response (operation) will be obtained. This process is called HTTP protocol [10]. The necessary code to obtain this control feature was created and dumped into the Node MCU with the help of the Arduino IDE. As soon as the Node MCU gets connected to the Internet, we can access the web server via the mobile app [14].

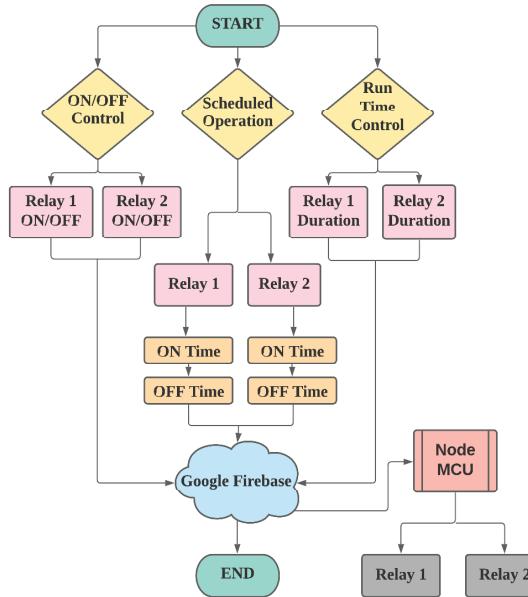


Fig. 3. Software Flowchart

Figure 3 represents the flowchart of the proposed mobile application. The three features are represented in the User Interface (UI). Initially, the Google firebase database is configured. The token and URL of the created database are included in the MIT app inventor. And the two variables are set in the database for relay 1 and relay 2. In on/off control, two buttons namely ON and OFF are provided. When the user clicks the button the status of the respective variables gets modified in the database [16,17]. The changes are retrieved back by the microcontroller and the relays are operated accordingly. In the run time control, the entered duration got saved and the device will be turned on for that prolonged duration. After the specified duration, the device gets turned off automatically. In the scheduled operation, the on-time and off-time can be configured by the user. For that respective duration, the device will get operated [12,13].

Figure 4 represents the charging duration (0 to 100 percent) of several mobile models for the respective battery capacity obtained from the field study and company website. With the help of this analysis, we can understand the performance of different mobile phones. Similarly, multiple plug-in devices will be operated / charged for a long duration. The proposed design also provides the customized

on-duration as user-friendly according to the model. This also prevents manual operation and aids more efficient control with the help of automation.

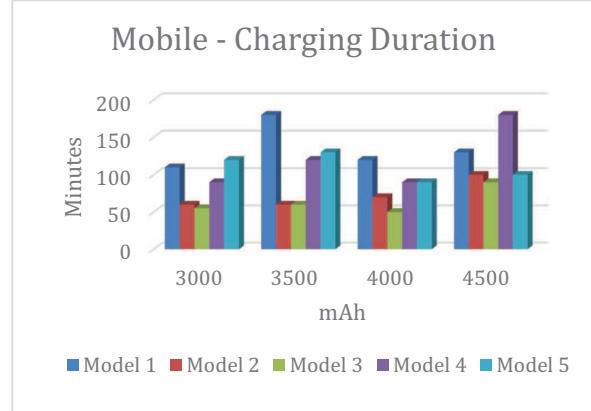


Fig.4. Analysis of Mobile Charging Duration

VI. HARDWARE AND SOFTWARE SETUP

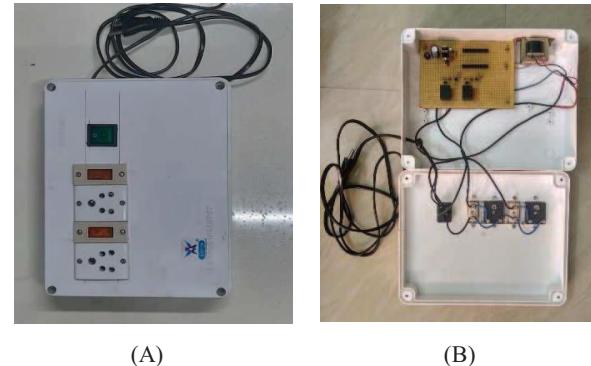


Fig. 5.A & 5.B Hardware part

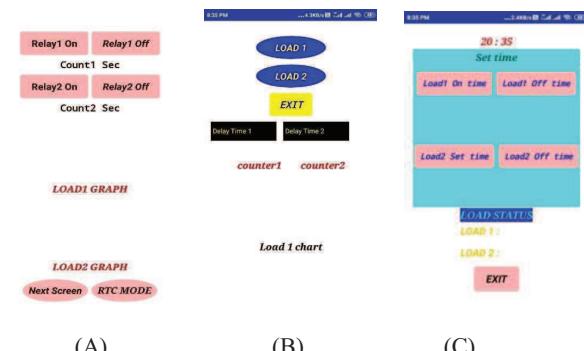


Fig.6.A, 6.B & 6.C Software part

This design proposes a model consisting of both hardware and software parts. The hardware part is designed like a switch box. The microcontroller and relays are included inside the hardware casing. The ends of the relay are connected to the switch ports. Two switch ports are provided in this model. As soon as the user clicks the option, the status of the assigned variable in the Google firebase is retrieved back by the Node MCU microcontroller. It is

programmed to act accordingly based on the input provided by the user. With the help of this design, the devices connected to the switch ports can be controlled easily via the Internet.

VII. RESULT AND DISCUSSION

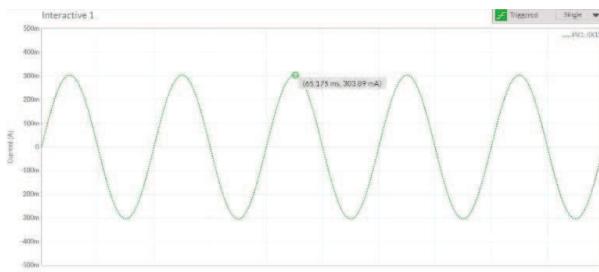


Fig.7. Load Current waveform

Figure 7 depicts the current waveform of the load (LCD TV) connected to our model. With the help of the proposed design, the user can control his devices based on his demand. Several control features can also be obtained. The hardware casing of the model provides complete insulation from the internal circuits. There are two switch ports provided in this design. Through that, we can connect two devices at a time. This design is suitable to control the small and large plug-in devices having a current rating below 5 A. The user interface of the software mobile application is more user-friendly which helps the users to understand it more easily. This mode will provide a huge benefit to the users by controlling their home appliances easily and efficiently. This proposed model will be less costly when compared with the conventional IoT devices in the market. It is easily portable and more user-friendly.

VIII. CONCLUSION

IoT devices have become inseparable in our daily life. This proposed design covers the drawbacks found in the earlier methods of control and has achieved the goal to the best. This model will help all kinds of people to control their home appliances easily from any part of the globe. Based on the overall results, we can confidently state that it will meet the needs of the users. The major value proportion of this design is a cost-effective and portable IoT home automation device. In the future, it can also be expanded to control high current rating plug-in devices.

REFERENCE

- [1] A.Karthi and G.Sabareesh, "Design of Internet Controlled Switch Box using IOT", Journal of Advances in Chemistry, 2016
- [2] Siddharth Karanchery and N. Rakesh, "Smart Power Socket using Internet of Things", Fifth International Conference on Inventive Computation Technologies (ICICT), 2020
- [3] Narasimha Swamy S and Kota Solomon Raju, "An Empirical Study on System Level Aspects of Internet of Things (IoT)", IEEE Access, 2020
- [4] S.Hrushikesava Raju, Dr.M.Nagabhushana Rao, N.Sudheer and P.Kavitharani, "IOT Based Home Automation System with Cloud Organizing", International Journal of Engineering & Technology, 2018
- [5] DInya Abdulahad Aziz, "Webserver Based Smart Monitoring System Using ESP8266 Node MCU Module", International Journal Of Scientific & Engineering Research, 2018
- [6] Prathmesh Shelke, Shubham Kulkarni, Swapnil Yelpale, Omkar Pawar, Ravdeep Singh, and Kirti Deshpande, "A NodeMCU Based Home Automation System", International Research Journal of Engineering and Technology (IRJET), 2018
- [7] Syed Kashan Ali Shah, Waqas Mahmood, "Smart Home Automation Using IOT and its Low-Cost Implementation", International Journal of Engineering and Manufacturing, 2020
- [8] Rampeesa Vijay, Thotakura Sainag, Vamsee Krishna A, "Smart Home Wireless Automation Technology using Arduino based on IoT", International Journal of Electronics & Communication Technology, 2017
- [9] Pooja More, Diksha Pawar, Gaurav Mahure, Sneha Bansod and Ankit Manke" Wi-Fi E Switch using Internet of Things", International Research Journal of Engineering and Technology (IRJET), 2018.
- [10] Wojciech Paweł Łasocha and Marcin Badurowicz, "Comparison of WebSocket and HTTP protocol performance", Journal of Computer Sciences Institute, 2021
- [11] Shaam Garg, Ayush Yadav, Shivam Jamloki, Aanchal Sadana, and Kusum Tharani, "IoT based home automation", Journal of Information and Optimization Sciences, 2020
- [12] Akram Khan, Abdullah Al-Zahrani, Safwan Al-Harbi, Soliman Al-Nashri, and Iqbal A.Khan, "Design of an IoT smart home system" published in IEEE, 2018.
- [13] O. Blanco-Novoa, T. Fernández-Caramés, P. Fraga-Lamas, and L. Castedo, "An Open-Source IoT Power Outlet System for Scheduling Appliance Operation Intervals Based on Real-Time Electricity Cost", IEEE International Conference on Internet of Things (iThings), 2017
- [14] Vishal Dineshkumar Soni, "Internet of Things based Energy Efficient Home Automation System", International Journal of Innovative Research in Science, Engineering and Technology, 2018
- [15] Chunnu Khawas and Pritam Shah, "Application of Firebase in Android App Development-A Study", International Journal of Computer Applications, 2018
- [16] Nilanjan Chatterjee, Souvik Chakraborty, Aakash Decosta, and Dr. Asoke Nath, "Real-time Communication Application Based on Android Using Google Firebase", International Journal of Advanced Research in Computer Science and Management Studies, 2018
- [17] Sergio Barrachina Mir and German Fabregat Llueca, "Introduction to Programming Using Mobile Phones and MIT App Inventor", IEEE Ibero-American Journal of Learning Technologies, 2020

Toxic Speech Classification using Machine Learning Algorithms

Pabba Sumanth
ECE, GRIET
Hyderabad, India
pabbasumanth1811@gmail.com

Syed Samiuddin
ECE, GRIET
Hyderabad, India
syedsamiu70@gmail.com

K. Jamal
ECE, GRIET
Hyderabad, India
kjamal24@gmail.com

Srikanth Domakonda
ECE, GRIET
Hyderabad, India
ssrikanth497@gmail.com

Pathi Shivani
ECE, GRIET
Hyderabad, India
shivanipathi2@gmail.com

Abstract— In today's era of online social media platforms, there has been a massive surge in the propagation of toxic content speech. They provide many betterments. However, persons with considerable differences in their viewpoints have contributed to an increase in lethality of people in internet posts and debates..With the outbreak of the pandemic, corporations, educational institutions, students, and the general public have all increased their usage in web sites. For a long time, the growing popularity of internet platforms like Twitter and Facebook has been a major cause of anxiety. These platforms not only allow for improved communication, but they also allow the users to express their thoughts, which are quickly shared with the rest of the world. Furthermore, given the diversity of these platforms' users' histories, beliefs, race, and customs, many of them choose to use disparaging, abusive, and antagonistic language while interacting with those who do not share their background. This online toxicity has been increasing exponentially by advancements provided by these social media platforms in this emerging world under the cloud of anonymity. Unlike manually, this problem can be solved using Machine Learning. Phrases like "Obscene", "Toxic", "Severe Toxic", "Threat", "Insult", "Identity Hate" are used mutually and hence have been incorporated under "Toxic" speech content. As a result, it is vital to recognise and eliminate toxic speech from internet - based social media networks naturally. The numerous varieties of Machine Learning approaches, such as traditional Machine Learning, ensemble approach are explored in this paper. We use a corpus collected from online platform twitter to do binary and multi-class classification and investigate two techniques.: (a) a method which consists in extracting of word embeddings and then generating the model; (b)Improving the existing models- RF, DT, VC, LR, KNN. Any other sort of social media comment can be analyzed using the proposed methods. By this, we developed a model that can classify given comments into different categories of toxicity with greater precision, recall, and accuracy score.

Keywords— Social Media Platforms, Machine Learning, Phrases, Ensemble Learning, Word embeddings, toxic and non-toxic classification.

I. INTRODUCTION

With the rapid growth of online social media sites such as Twitter, Facebook, Instagram, Youtube, and Snapchat in today's age of internet connections, more than half of the population now has access to the internet. The world's population wants to communicate and connect over the internet, these websites being able to connect with a large group is a valuable skill by creating and sharing content to interact with a wide audience.

People's inability to comprehend and recognise the fact that opinions, ideas, and views of individuals from different genders, linguistic backgrounds, and traditions is frequently driving the force behind the spread of this toxicity, and thus hate content is directed at a particular sex, religious doctrine, demographic origin, or race based society. Adolescents make up a large portion of the population targeted for hatred. As a result, a large portion of our dataset consists of posts and comments circulating among teenagers.

According to public research studies, over 25% of respondents feel insecure in their communities as a result of the spread of online hate, and more than 80% are attempting to monitor and counteract cyberbullying. In order to preserve the standard of their content, platforms such as Facebook and YouTube have to erase 3 million posts and 25,000 videos in 2018. The proliferation of such information is leading to an increase in violence, such as communal riots and lynchings, throughout the world. Furthermore, as a result of the toxicity on the internet, children under the age of 25 are twice as likely to participate in suicide conduct. Two worldwide organisations attempting to gain a deeper understanding, distribution, and prevention of online hate speech are the United Nations Human Rights Council and the Online Hate Prevention Institute..

The prevalent use of various conversation languages and short forms, such as "wtf", "idk", "tbh", "asap", etc., emojis are used to make information more effective and evocative, and numerous repeats of those are used to underline emotions in a remark, such as "Fuck yourself! Get your fucked up on! ", researchers and authorities will find it extremely difficult to

analyse the material and determine whether it is infested as a result of all of this, is used as an indication of the sender's level of contempt for the addressee, and the massive use of grammatical symbols, such as "Shut up!!!! You bloody!!!",

Hate speech is a premeditated tactic of prejudice, violence, and hatred that lacks an international legal designation. A toxic comment is one that is nasty, insulting, or unreasonable, and is likely to turn other users off from a discourse. A step of the algorithm of sentiment analysis is toxic comment categorization. When you say something is obscene, you're expressing that it offends you because it contains sex or violence in an inappropriate and frightening manner. The terms are frequently misunderstood, as the three examples in Table 1 demonstrate.

TABLE 1. Examples of comments

S.No	Comments	Classification
1	"unblock me or i'll get my lawyers on to you for blocking my constitutional right to free speech"	Toxic
2	Tony Sidaway is obviously a fistfuckee. He loves an arm up his ass.	Obscene
3	You are so retarded	Insult

II. RELATED WORK

Regardless of the fact that toxic speech identification is a considerably new field related to research, it has achieved a number of significant advancements. There has been study on the social networking sites Twitter[1], Facebook[5], YouTube[2], and Reddit[15]. Fortune and Nunes[3] investigated the motives for doing hate speech detection research and offered a comprehensive review of the field's future possibilities. In [18], Waseem employed a dataset of 16K tweets to categorize them as gender stereotypes, racial prejudice, or none of the above. He did the best, using the LR algorithm when compared to other techniques like character and word n-grams. Gaydhani et al. [4] used a mix of three datasets to perform logistic regression. She discovered that using logistic regression and a term frequency and inverse document frequency vectorizer resulted in a 95.6 percent accuracy. In [1], Davidson, he analyzed a 24k tweet corpus that he divided into three categories: hate, offensive, and neither. He then used NLP techniques on the tweets, such as extracting the base form of word, text clustering, term frequency vectorization, and sentimental techniques of vader, before running various supervised learning algorithms, the best of which was LR with L-2 regularization. However, they determined that using lexical approaches, it was impossible to discern between Hate and Offensive material.

Machine learning algorithms have substantially aided hate speech detection and social media content analysis in general [8]. Offensive words such as TS and online bullying have been the most researched issues in NLP during the last two decades. Machine learning algorithms have shown to be highly useful in terms of social media data analysis for the finding and categorization of offensive statements [13]. Machine learning algorithm research advances have had a significant impact on a range of areas, resulting in several crucial tools and models for evaluating vast volumes of data in real-world settings, such as social media network content analysis [9].

The survey's authors [19] provided a brief assessment of eight hate speech detection strategies and approaches. From among eight ways are vectorization, vocabularies, N-gram, sentimental analysis, pattern-based approach, grammatical features, rule-based and bag of the word approach. The research's disadvantage seems to be that techniques such as the ensemble approach were not considered.

Hate speech posters typically target people based on their religion, race, political affiliation, gender, marital status, ethnicity, health condition, handicap, and nationality [10]. The amount of data created by SM sites is expanding geometrically every day, which is referred to as big data [8]. As indicated in figure 1, the world's population is around 7.7 billion [12], and the following estimated population is actively linked on one or more social sites [14].

It is critical to study with such a vast group and to comprehend human behavior trends. A situation like this, which can be produced by a vast population, cannot be overlooked.

III. PROPOSED METHODOLOGY

The purpose of the project is to put a set of comments into one of six categories, which are:

- **Toxic:** A toxic comment is one that is unpleasant, disrespectful, or irrational and is likely to drive other users away from a conversation. Toxic comment classification is a subtask of sentiment analysis.
- **Severe Toxic :** Adverse effects that arise following the repeated or continuous administration of a test sample for a significant portion of one's life span are referred to as Severe Toxic.
- **Obscene:** When you say something is obscene, you're implying that it offends you because it involves sex or violence in a way that you find offensive and disturbing.
- **Insult:** An insult is a purposefully rude action or manner of speech, as well as a lack of regard, esteem, or courteous behavior.
- **Threat:** It refers to a threatening message in a terrifying manner or a threatening message in a frightening manner.

A. Toxic Speech

Toxic speech is any conversation, including speaking, literature, and conduct, that criticises or defames an individual

or a team based on private data or hiding characteristics [16]. Religious, ethnic background, citizenship, civil status, general health, racism, appearance, disability, gender identity, descent, sex, or other identifying characteristics are all protected characteristics. The majority of hate speech messages on SM are composed of textual, toxic speeches, on the other hand, are disseminated via visuals and noises [11]. As a result, text categorization is the best chance for any attempt to handle this problem from a computer standpoint.

A simpler and more exact definition of hate speech has been shown to simplify annotators' jobs and, as a result, enhance the rate of agreement amongst annotators [9]. Apart from toxic speech, the other harmful online behaviors that need to be addressed, such as online bullying. This online bullying is also called Cyberbullying. Online bullying, a type of social harassment [17], is defined as "repeated aggressive behavior through social media in an attempt to threaten or harm persons who are unable to defend themselves readily" and is particularly prevalent among teenagers .

Cyberbullying, cyber-hate, and toxic speech are all examples of extremely toxic online behavior [7]. When a victim's sensitive or protected trait is the object of an attack, online bullying can be considered hate speech. Hate speech differs from cyberbullying in that it affects more than just one individual and has ramifications for the entire group or society.

B. Natural Language Processing (NLP):

This is a branch of artificial intelligence that deals with processing human-readable language and making machines comprehend it. The primary goal of NLP is to extract, analyse, and comprehend linguistic forms in desire to make understanding of them. To evaluate information and extract relevant meaningful data, the bulk of NLP approaches depend on ML algorithms. NLP applications include sentiment analysis, personal assistant programmes like Google, Siri, Alexa, and others, language translators, and so on.

There are two components of NLP. They are:

1. Natural Language Understanding (NLU):

Natural Language Understanding (NLU) extracts metadata from material such as concepts, entities, keywords, emotion, relations, and semantic roles to assist machines comprehend and analyse human language. NLU is mostly used in business applications to comprehend the problem of a client in both spoken and written language.

2. Natural Language Generation (NLG):

Natural Language Generation (NLG) is an interpreter that converts electronic information into basic language. Textual organization, Syntax preparing, and Word Synthesis are critical components.

These would be accomplished by employing a variety of programs to determine the semantics of each phrase in a paragraph and gather relevant data. Natural languages incorporate words with emotive connotations. When employed in different settings, the same statement might have distinct meanings. As a result, robots must be able to fathom and

evaluate the contents of a statement in light of its frame of reference.

NLP may be used to process text in a variety of ways.

Some frequent NLP stages are as follows:

1. *Linguistic Analysis*
2. *Phonetic Analysis*
3. *Syntactical Evaluation*
4. *Inclusion of Discussion and debate*

C. Machine Learning

The predictive approach is another name for this concept. This procedure is based on a dataset that has been coded with or without guidance and may be utilised for preparing the model. This labelled dataset is used to train the model by supervised or unsupervised learning methods in order to categorise the comments into different levels.

The following algorithms are part of this study:

Logistic Regression (LR): Regression analysis is conducted when the dependent variable i.e. output data is categorical or binary. These are four types of logistic regression and mainly we use binary LR. Here, there are three link functions such as logit, normit (probit), and gompit.

Support Vector Machine (SVM): This procedure is commonly used in high-dimensional spaces for classification, regression, and outlier detection. It categorises sets of data by forming an N-dimensional hyperplane (representing N characteristics).

Decision Tree (DT): One of the most complete machine learning methods is the Decision Tree (DT). Decision trees, which aid in selecting the most significant aspects of a data collection, may not readily be captured as linear patterns. Factors like entropy and information gain influence how a decision tree splits.

KNN Algorithm: It is abbreviated as k-nearest neighbor. In this, the algorithm first learns from the training data and makes different groups by classifying those corpus. furtherly, it classifies the new data received into those classified groups thus making the model more efficient by converting into understandable format. This means that employing the K-NN methodology, new information may be efficiently categorized into a definite category.

D. Ensemble Learning

The ensemble method includes nothing more than drawing on collective wisdom. To put it another way, the best single classifier is indeed greater than the competition of numerous classifiers' predictions [20]. This grouping technique was developed to address the shortcomings of several machine learning algorithms while increasing their strengths. Each model, of course, has shortcomings of its own; so, no model is perfect. Though ensemble techniques aim to combine the benefits of several models to produce a better efficiency than any particular approach.

Random forest, bagging methodology, and boosting method are some of the numerous forms of ensemble approaches. Each of these approaches has advantages and disadvantages when it comes to dealing with hate speech.

Random Forest (RF): It is a supervised learning algorithm that is employed to learn from the training data and predict the output for the test data. This approach generates a clustering model with more accuracy and less overfitting cases of each feature. It is the most used algorithm and employed in every ML related real-time problem. It contains a number of decision trees from the subsets of the datasets and takes the average to improve the model accuracy. It also supports larger datasets with great dimensionalities.

Voting Classifier (VC): A Voting Classifier is a predictive learning model that learns from other models and evaluates an output based on the efficiencies of the result obtained from those models. Generally, it combines the results of each ML algorithm passed into VC and predicts the output based on the trained models. Instead of using various algorithms separately, their combination brings a very fruitful output with efficient values. Its efficiency may be less than the other classifying models, but this is the most valid one.

IV. EXPERIMENTAL SETUP

Fig.1 shows the essential concept of the approach suggested in our research study for interpreting text into one of the following categories: toxic, severe toxic, obscene, insult, threat, identity hate. As a first step we created a dataset using social media comments for the research. The rest of the study involves the use of various ML classifiers.

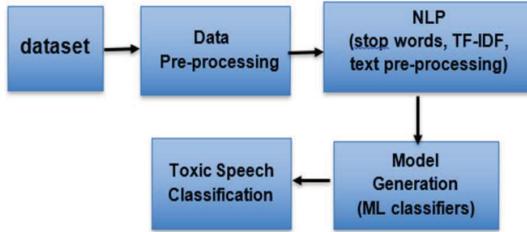


FIGURE 1. Flow chart of toxic speech classification

A. Data Set

Our research was based on a set of data taken from the kaggle. This dataset is created from youth's conversation from social media platforms. It groups these conversations into the following categories as shown in table-2. It consists of nearly 160k tweets. This expanded the amount of our algorithm's training dataset and, more importantly, it fixed the problem of unbalanced labels in the dataset.

TABLE 2. Statistics of Comments

Classification	Toxic	Severe Toxic	Obscene	Threat	Insult	Identity Hate
Count	15294	1595	8449	478	7877	1408

For this research we used one three-fourth of the corpus to train the model and the rest to test.

B. Pre-Processing

Information is relatively raw and unstructured in the real world, and it must first be organised before being treated to eliminate unwanted information, padding the incomplete data, reduce inconsistencies, and increase the corpus's quality and processing capabilities. NLTK, a python supported library, was used to perform purification procedures in the given dataset.

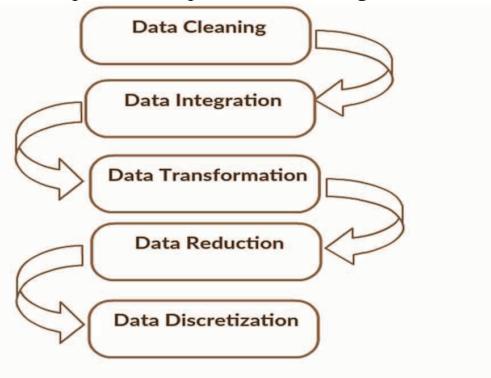


FIGURE 2. Steps in preprocessing

For each phrase in the dataset, we applied the below cleaning techniques as shown in fig.2.

- The short phrases were broadened to their original forms, such as "don't" changes to "do not," "can't" becomes "can not," and so on.
- All stop words have been deleted from those texts, leaving just significant words. We discovered that some of the keywords in the corpus might be classed as usually repeated words, but they weren't in the NLTK words vocabulary, so we created a related repeat words lexicon to remove them out of the data.
- Because, corpus included comments, there were numerous hyper-links, profile mentions, and mentions that needed to be eliminated because they were useless for text categorization. To do this, we utilised the Python package "re," which is based on regular expressions.
- Following that, we lemmatize those frequent words in diverse forms to their root word, removing repetition and guaranteeing that the same word in several forms is not considered separate. To eliminate ambiguity in words, for example, tagging portions of speech is sometimes used. The term "bear" will be employed as a different part of speech depending on the context. As a result, if POS tagging is not utilised, the content may be misconstrued. We also converted each word to lowercase so that the similar sentence may not be treated differently during word vectorization due to character differences, as they are syntactically correct.
- Finally, all punctuation and whitespaces were deleted from the set.

Data preprocessing is the process of converting and representing data in a manner that is suitable for processing.

Using a variety of data visualisation approaches, preprocessed data is shown. It is amongst the most important methods for analyzing observational comprehension. It might come in handy while looking at a dataset and extracting insights from it, and also recognising patterns, faulty information, anomalies, and related issues.

C. ML Models

From Fig.1, after preprocessing our next step is to train the model using different algorithms present in the ML library. In this research we analysed the data using Logistic Regression as a base algorithm. Following that, we employed a Support Vector Machine algorithm for better training the model. In continuation, we also used algorithms related to classification and regression. Finally, we created a Voting classifier that combines many machine learning models to get superior results. We observed that these methods are inefficient at identifying clusters or series of phrases which are critical for distinguishing between the different categories of toxicity.

Further, we used clusters using models in order to improve the accuracy. We used Random Forest by keeping estimators as default. Among all the methodologies, the best result obtained by using Random Forest. Scores obtained from each of the algorithms are recorded in the table-3.

D. Key Indicators for Toxic Speech Classification

Performative evaluation is a multidisciplinary study topic that is frequently achieved via the use of performance evaluation measures. These evaluation metrics are conceptual constructs derived by subtracting observed and anticipated values [21]. To evaluate the efficiency of hazardous speech detection algorithms, the standard accuracy, reliability, and F-measure measurements are usually utilised. They are mostly used due to the unequal structure of the toxic speech dataset. For any balanced dataset, accuracy is the best option. All major metrics are explained in depth in [8] and [6].

Assume our model was trained to distinguish between toxic comments and non-toxic comments in tweets. We considered a group of 25 comments with 10 comments classified as toxic speech and 15 classified as non-toxic speech. 11 tweets were identified as toxic speech by the model. 8 of the 11 tweets found were indeed hate speech (true positives), whereas the other 3 were not (false positive). The programme incorrectly categorised 4 tweets as hate speech (false negative), whereas 11 tweets were correctly removed as non-hate speech (true negative).

1. PRECISION (P_r):

Precision is described as the ratio of True Positives to all positives. Precision was used by the following researchers to assess the performance of their models.

$$\text{Precision} (P_r) = \frac{TP}{TP+FP}$$

Here TP is abbreviated as true positive. From the above scenario, the value of TP is 8. False positive is shortly represented as FP. From the above, FP is 3. So, precision is $8/10 = 0.8$.

2. RECALL (R_c):

Out of all correct positive examples in the dataset, recall is the number of correct positive class predictions produced. Many papers used recall to evaluate their findings.

Mathematically,

$$\text{Recall} (R_c) = \frac{TP}{TP+FN}$$

Here FN is abbreviated as false negatives. From the scenario above, the value of FN is 4. The true negative is shortly represented as TN. The value of TN from above is 11. So we get recall is $8/12 = 0.6$. This represents that the model is able to identify 60% toxic comments correctly.

3. F-MEASURE:

F-Measure generates a single score that combines both accuracy and recall problems in a single number. The weighted harmonic mean (whm) of accuracy and recall is the F-measure (F) or F1-score (F). When the dataset is imbalanced, this assessment metric is often used. It was used to assess the effectiveness of hate speech prediction models in various studies.

Mathematically,

$$\text{F-Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

4. ACCURACY (A_r):

The ratio of correct predictions to total observations is known as accuracy. For the two-class problem, the accuracy of a model is greatest if and only if we have a symmetric corpus in which the values of False Positive (FP) and False Negative (FN) are nearly identical. In many and imbalanced data sets, accuracy is not the best option; therefore, other evaluation measures, such as the F1-score, may be evaluated. It was used in many researches. Accuracy (A) can be stated mathematically as:

$$\text{Accuracy}(A) = \frac{TP+TN}{TP+FP+FN+TN}$$

TABLE 3. OUTCOMES

	LR	SVM	DT	RF	KNN	VC
Precision	0.90	0.91	0.85	0.92	0.95	0.94
Recall	0.82	0.81	0.78	0.76	0.09	0.73
F1-Score	0.86	0.86	0.81	0.83	0.16	0.83
Accuracy	0.86	0.86	0.81	0.84	0.53	0.83

V. RESULTS

We performed various algorithms including Space vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K Neighbors (KNN), Logistic Regression (LR), and Voting classifiers (VC), with SVM surpassing the others. We expected Random Forest to perform the best among the others, due to the fact that RF is being used widely for classification tasks. But in spite of that in our case, SVM performed slightly better, demonstrating that it can locate a small set of essential words and consider them while adjusting the noise in them. In Table, we have mentioned the precision, recall, F1-score, and accuracy of each algorithm. The accuracy of each algorithm is shown in figure-4.

1.GET FUCKED UP. GET FUCKEEED UP. GOT A DRINK THAT YOU CANT PUT DOWN???/ GET FUCK UP GET FUCKED UP. I'M FUCKED UP RIGHT NOW!!!!!!.

Output:

Toxic Detection

Enter a comment to check if it is toxic!

GET FUCKED UP. GET FUCKEEED UP. GOT A DRINK THAT YOU CANT PUT DOWN???/ GET FUCK UP GET FUCKED UP. I'M FUCKED UP RIGHT NOW!!!!!!.

Predict

Toxic: 1.0

Severe Toxic: 0.31

Obscene: 0.99

Insult: 0.38

Threat: 0.04

Identity Hate: 0.3

FIGURE 3. Final Output for a comment

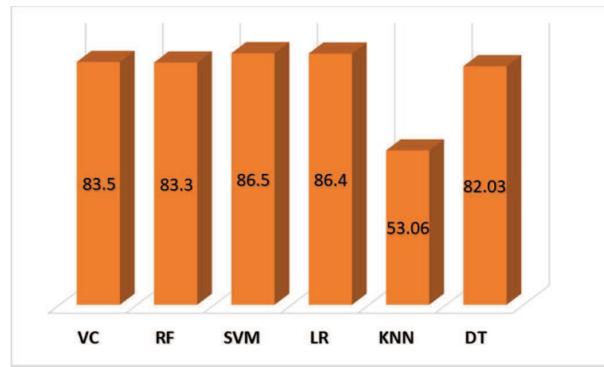


FIGURE 4: Accuracy of ML algorithms

VI. CONCLUSIONS AND FUTURE SCOPE

In this research, we mainly concentrated on the comments that are common among youth's conversation in social media platforms. We created models using various algorithms of Random forest, KNN, etc. We are able to classify the comments into six different categories, among all algorithms Support vector machine and Logistic regression gave more accurate results. The model not only categorizes a given sentence as toxic or non-toxic, but also provides percentages of Obscene, Toxic, Severe Toxic, Hate, Threat, Identity Hate. On the trained model, we observed 0.91 precision, 0.81 recall, and 86.5 accuracy.

In the future, we plan to test and examine different models and paradigms, such as convolutional neural networks coupled with recurrent neural networks and feature extraction methods used for hazardous speech identification. We believe that the cascading model is a good concept. Despite the fact that we just built a simple version of the model, a more complex version might increase the model's efficiency and accuracy.

REFERENCES

- [1] T. Davidson, D. Warmsley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," arXiv, no. Icwsm, pp. 512–515, 2017.
- [2] P. L. Teh, C. Bin Cheng, and W. M. Chee, "Identifying and categorising profane words in hate speech," ACM Int. Conf. Proceeding Ser., pp. 65–69, 2018.
- [3] P. Fortuna and S. Nunes, "A survey on automatic detection of hate speech in text," ACM Comput. Surv., vol. 51, no. 4, 2018.
- [4] A. Gaydhani, V. Doma, S. Kendre, and L. Bhagwat, "Detecting hate speech and offensive language on twitter using machine learning: An N-gram and TF IDF based approach," arXiv, 2018.
- [5] F. Del Vigna, A. Cimino, F. Dell'Orletta, M. Petrocchi, and M. Tesconi, "Hate me, hate me not: Hate speech detection on Facebook," CEUR Workshop Proc., vol. 1816, pp. 86–95, 2017.
- [6] V. Balakrishnan, S. Khan, T. Fernandez, and H. R. Arabnia, "Cyberbullying detection on Twitter using big five and dark triad features," Personality Individual Differences, vol. 141, pp. 252–257, Apr. 2019.
- [7] R. Alshalan and H. Al-Khalifa, "A deep learning approach for automatic hate speech detection in the Saudi Twittersphere," Appl. Sci., vol. 10, no. 23, pp. 1–16, 2020.

[8] M. A. Al-Garadi, M. R. Hussain, N. Khan, G. Murtaza, H. F. Nweke, I. Ali, G. Mujtaba, H. Chiroma, H. A. Khattak, and A. Gani, “Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges,” IEEE Access, vol. 7, pp. 70701–70718, 2019.

[9] B. Ross, M. Rist, G. Carbonell, B. Cabrera, N. Kurowsky, and M. Wojatzki, “Measuring the reliability of hate speech annotations: The case of the European refugee crisis,” 2017.

[10] T. Granskogen and J. A. Gulla, “Fake news detection: Network data from social media used to predict fakes,” in Proc. CEUR Workshop, vol. 2041, no. 1, 2017, pp. 59–66.

[11] C. Ring, “Hate speech IN social media: An exploration of the problem and its proposed solutions,” 2013.

[12] L. Tamburino, G. Bravo, Y. Clough, and K. A. Nicholas, “From population to production: 50 years of scientific literature on how to feed the world,” Global Food Secur., vol. 24, Mar. 2020, Art. no. 100346.

[13] G. Weir, K. Owoeye, A. Oberacker, and H. Alshahrani, “Cloud-based textual analysis as a basis for document classification,” in Proc. Int. Conf. High Perform. Comput. Simul. (HPCS), Jul. 2018, pp. 629–633.2020, no. Icwsrm, pp. 452–463, 2020.

[14] S. Paul, J. I. Joy, S. Sarker, A.-A.-H. Shakib, S. Ahmed, and A. K. Das, “Fake news detection in social media using blockchain,” in Proc. 7th Int. Conf. Smart Comput. Commun. (ICSCC), Jun. 2019, pp. 1–5.

[15] A. Mittos, S. Zannettou, J. Blackburn, and E. De Cristofaro, “And we will fight for our race! a measurement study of genetic testing conversations on reddit and 4chan,” Proc. 14th Int. AAAI Conf. Web Soc. Media, ICWSM

[16] J. van Dijck, “Governing digital societies: Private platforms, public values,” Comput. Law Secur. Rev., vol. 36, Apr. 2019, Art. no. 105377.

[17] P. Smith, J. Mahdavi, M. Carvalho, and N. Tippett, “An investigation into cyberbullying, its forms, awareness and impact, and the relationship between age and gender in cyberbullying,” Res. Brief, London, U.K., Tech. Rep. RBX03-06, Jul. 2006, pp. 1–69.

[18] Z. Waseem, J. Thorne, and J. Bingel, “Bridging the Gaps: Multi Task Learning for Domain Transfer of Hate Speech Detection,” Springer International Publishing, 2018.

[19] A. Alrehili, “Automatic hate speech detection on social media: A brief survey,” in Proc. IEEE/ACS 16th Int. Conf. Comput. Syst. Appl. (AICCSA), Nov. 2019, pp. 1–6.

[20] A. Géron, Hands-On Machine Learning With Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. CA, USA: O’Reilly Media, 2017.

[21] A. Botchkarev, “New typology design of performance metrics to measure errors in machine learning regression algorithms,” Interdiscipl. J. Inf., Knowl. Manage., vol. 14, no. 113, pp. 13–21, 2019.

[22] A. Botchkarev, “New typology design of performance metrics to measure errors in machine learning regression algorithms,” Interdiscipl. J. Inf., Knowl. Manage., vol. 14, no. 113, pp. 13–21, 2019.

Implementation of ANN for Predication of Performance and Emission Characteristics in CI Engine

1st Dr.Narendranathan.S.K
Mechanical Engineering
Agni College of Technology
Chennai, India
sknarengopi@gmail.com

2nd Dr.M.Thiyagu
Mechanical Engineering
Saveetha School of Engineering
Chennai, India
thiyagu1080@gmail.com

3rd Mr.Shankar.A
Aerospace Engineering
Agni College of Technology
Chennai, India
shanky977@gmail.com

4th Mr.T.N. Sureshkumar
Mechanical Engineering
DhaanishAhmed College of Engineering
Chennai, India
tn_sureshkumar@yahoo.co.in

Abstract— Over the past few years, Global warming and huge Depletion of Petroleum resources have increased the need for alternate fuels for Internal Combustion (IC) Engine. In recent years, bio-diesel extracted from vegetable oil is becoming popular since it is renewable and environment friendly. There are multiple parameters involved in determining the performance of IC engines and these parameters are linked to each other. Since multiple inter-related factors are involved, optimizing the performance of IC engines is a challenging task. Existing methods show poor performance and also they are time consuming. Due to the ability to analyze complex and huge volume of data, Applications of Machine Learning and Deep Learning algorithms is increasingly used in various fields. Artificial Neural Network (ANN) is used in the field of optimizing the performance of IC engines. In this paper, recent developments have been analyzed and the use of ANN in predicting various parameters is discussed. Three different blends of diesel, jatropha oil and neem oil was taken and the input parameter load is varied to predict the output parameters. Experiments shows that ANN has produced better optimization.

Keywords—Artificial Neural Network, Bio-Fuels, Internal Combustion Engines.

I. INTRODUCTION

Fuel Energy[1] plays a vital role in day-to-day life and it is an important factor in deciding the economy and to increase the standard of living. Transportation is the heart of every sector and the rise in fuel price will lead to the increase in price of all commodities and thus affects the life of common people. Use of fuel is increasing in a rapid rate which in turn leads to the depletion of petroleum fuels which is non-renewable. Also the use of vehicles with petroleum fuels leads to environmental pollution and increases the effects of global warming. The quantity of carbon dioxide which is released into the environment has gone beyond the permissible amount and has contributed a lot to the greenhouse effect.

Abnormal circumstances like increased pollution in air, increased acidic level in air has contributed to greenhouse effects. So it is a crucial situation to find a renewable energy source as an alternate to petroleum fuels. Many of the things around us can be converted into energy if proper methodology is used. Many researchers have found alternate fuels and each has its own advantages and disadvantages.

Oil crops[2] are used to produce bio-fuels and it will be a better alternative to petroleum based fuels since it is both renewable as well as environment friendly. Seeds such as

jatropha, sunflower, palm, flax, punnai, soya bean, lemon grass etc can be used to extract bio-diesel. Bio-fuels can be extracted both from edible as well as non-edible oils. Initially bio-fuels was extracted form edible vegetable oils. Since it affects the supply of food to common people, extraction of bio-fuels from non-edible oil was found. Bio-Fuels have properties that are similar to petroleum based fuels. Engines have to be slightly modified to adopt to unique characteristics of bio-fuels.

Currently, the use of bio-fuels is less but in future there will be an increased level of use due to its preferred characteristics. Bio-diesel can also be extracted from fats of animals like pork, fish, beef, duck etc., which is also a renewable form of energy and environment friendly. Good volume of bio-fuel can be extracted from the wastes that are thrown from fish. A huge volume of wastes are produced every year from the food industries that process fishes. There are totally 38 processing units for wastages from food wastes in India. Experiments have been done for measuring the performance of combustion engines using fish based bio-fuel. It shows almost same performance as petroleum based fuels but with less emission of carbon-di-oxide.

To bring the unique characteristics of petroleum[3] based fuels, several process such as esterification, blending with other oils etc., are carried out. In some cases, the extracted bio-fuels is found to be less denser than the normal petroleum based fuels. Some researchers have tried to mix vegetable oil with petroleum based fuels and found that these type of oils produced good performance characteristics. Also blending bio-fuels with petroleum based fuels does not require major changes to the existing engine structure.

Bio-fuel extracted from vegetable oils prove to be a good alternative. But care has to be taken that when more bio-fuels are extracted it should not increase food scarcity for common people. Hydrogen can be extracted from water and it can also be used as an alternative to petroleum based fuels. They show performance characteristics similar to that of bio-fuels. Hydrogen is found to be good in working with Spark Ignition(SI) engine. But when used with Internal Combustion (IC) engines, there is a need for ignition source in order to start with the process of combustion. So, dual mode of operation is

preferable and diesel can be used to start the combustion process.

The parameters that affects the performance of engines[4] should be carefully noted, since it needs to be observed in bio-fuels. Many studies focus on finding an alternate to the existing petroleum based fuels. But care has to be taken in the amount of modifications that are required in the present diesel engines for using the alternate fuels. Little modification is acceptable but changing the entire design of the engine is not acceptable since it involves huge cost.

Complex computations are needed for finding the performance of engines with different blend of oils and at various operating systems. Traditional methods for capturing the engine performance and optimization are time consuming and comes at high cost. Recently, computer based Artificial Intelligence (AI) methods are gaining its popularity because of its ability to capture the complex nature. AI based methods are almost used in all fields including medicine.

Artificial neural network is an AI based method that tries to resemble the way in which the human brain works. Our human brain is able to find the complex relationship between elements by processing large data in less time. Since, Artificial Neural Networks (an AI method) are best at capturing the complex relationships, They can be suitable for measuring the performance of engines under various circumstances. Genetic Algorithms can be combined with ANN to bring out better performance. Several algorithms like fuzzy logic are increasingly used in combination with ANN methods.

When the required amount of input data is given, ANN is able to predict the needed output parameters precisely. The input data will be generated by conducting experiments in real time with various blends of oils. ANN can give the needed output with less resource consumption. ANN is gaining its popularity due to its self learning capability and its ability to compute non-linear computations. In this paper, various ANN based methods that are used in predicting the engine behavior was studied and the parameters that are needed for the better optimization of the output engine was found and listed. The various bio-fuels and their needed level of modification in the existing engine is also studied.

II. RELATED WORKS

In [5], six different sets of bio-fuels and petroleum fuels blends (in ratio of 10% and 90%) are used to find the combination that provides better results. The thermal behavioral properties as well as the physico-chemical properties of bio-fuels are analyzed inorder to replace them in place of petroleum based fuels. There is a performance almost equal to petroleum based fuels and a reduction in CO₂ emission is also observed.

In [6], the authors investigated the performance of IC engines by using calophyllum inophyllum methyl esters blend. Adaptive Neuro Fuzzy Inference System (ANFIS) is used to

model and to do experimentation. The experiments are conducted by various inputs such as different blends of bio-fuels, different cylinder pressure, different heat release by varying loads etc., The output parameters that are observed includes emission of carbon monoxide and smoke and also Break Specific Fuel Consumption (BSFC). Trial is made to bring down the errors by using the output performance from ANFIS.

In [7], fish based bio diesel blend is taken for evaluating the performance of Compression Ignition (CI) engine. Fuzzy inference System (FIS) is used to find the performance of CI engines and to study their emission characteristics. The input parameter given is same injection speed with varying loads. Experimental results shows that there is increased accuracy and decreased time consumption on using FIS system for modeling.

In [8], the bio-fuels extracted from karanja oil and reselle oil are used in combination with petroleum based fuels with varying blend proportion. The ratio of compression as well the load given to engine is varied to carry out the experiments. Artificial Neural Network is used to evaluate the performance of engines and the experiments shows improved performance of output parameters.

In [9], authors have experimented to use Hydrogen and oil extracted from Lemon grass as a bio-fuel. Experiments were initially done with petroleum fuel and then different proportions of Hydrogen and Lemon grass oil is used in dual-fuel mode. The ANN with back-propagation is used find the dependency between the output parameters and varying input parameters. This method tries to improve output parameters by optimizing the input parameters given to the engine.

In [10], ANN is used for optimization of output parameters and experiments are conducted with varying learning rates. Genetic Algorithms are used along with ANN to further improve the performance. In [11], the bio-fuel extracted from Jatropha oil is used as an alternate to petroleum based fuel. The optimization based on Taguchi's method is used to enhance the performance of combustion engines. 10 parameters that plays a crucial role in determining the performance of engines was used. The proposed method correctly identifies the parameters that plays a vital role in optimizing the engine performance.

In [12], the bio-fuel extracted form bovine manure through the process of anaerobic fermentation is used for experimentation. This method tries to improve the Thermal efficiency and volumetric efficiency in a spark ignition engine. Variations of methane concentration and different load values of engines are given as input. ANN is used to find the optimization of the output parameters. The results produced by ANN produces higher correlation among the input and output parameters.

In [13], the bio-fuel extracted from non-edible honne oil is used in combination with petroleum based fuels. Tests were done by varying the input parameters like ratio of

compression, load, pressure of injection etc., Minitab is used to develop a model as a function of the input parameters and it tries to optimize the engine performance. ANN is used in combination with Genetic algorithm to predict the optimization of engines based on input data values. The multilayer neural network is used along with the concept of back propagation.

In [14], hydrogen has been used as an alternate to petroleum based fuels and the burning capability of hydrogen fuel is able to produce less CO₂ emission. Hydrogen can be used along with petroleum fuels in dual-fuel mode as well. ANN is used for modelling and to predict the optimization of output parameters like BSEC, CO₂ emission, NOx etc., Fuzzy concepts are used along with ANN and both logarithmic sigmoid and hyperbolic tangent sigmoid are also used.

In [15], oxygenated fuel is used to evaluate the performance of engines. ANN is used for modelling and it is able to find the relationship between the input and output parameters. This method is able to find the favorable engine operating condition with different blends of fuel. In [16], various blends of fish oil with petroleum based fuels was tried and it uses the concepts of Fuzzy logic as well as genetic algorithm to find the optimum blend ratio. The load of engine was varied while keeping the engine speed as constant. This method is able to produce low error rates.

In [17], the experiments were conducted with varying blends of palm based bio-fuel, petroleum fuel and ethanol. ANN is used to model the engine with varying load of engine and maintaining constant engine speed. From experiments, it is found that the logistic sigmoid activation function in combination with the back propagation algorithm gives better optimization and higher degree of accuracy.

In [18], experiments are carried out in dual fuel mode using hydrogen and bio-fuel extracted from Jatropha. Meager experimentation is carried out and the data needed for training ANN has been obtained. If sufficient quantity of data is obtained, ANN will give increased accuracy. Seven different algorithms with varying transfer learning function has been applied and a comparison of their performance was done. The Back propagation algorithm in combination with logarithmic sigmoid function produced better performance than the other combinations considered.

In [19], back propagation algorithm in ANN is used for evaluating the model. The ANN network is used for finding the relationship between the input parameters and output parameters. Experiments were done using palm oil and waste cooking oil based bio-fuels. Multilayer perceptron based ANN is developed. Experiments shows that ANN can model the performance of IC engines in a better way.

In [20], ANN was used to model the operating factors of IC engines and also to optimize the working of IC engines. Setting up the experimental design is an important parameter in deciding the quality of output produced. Here, weights have been adjusted and learning functions have

been varied to assess the performance of the ANN model. From experiments, it is found that the ANN with backpropagation performs better than the other modeling algorithms. The experiments shows accuracy of above 95%.

III. Methodology

Artificial Intelligence (AI)[21] is an emerging field which aims at producing machines that have intelligence similar to human beings i.e. machines will be able to think and act like human being. The field of Artificial Intelligence is gaining popularity in almost all fields due to the migration towards automation. AI machines helps mankind to save time, energy and money and can be able to bring out results quickly.

Machine Learning (ML) is a sub-division of AI which allows the machines to learn from their previous experience and it uses statistical based algorithms. The difference between AI and ML is that AI machines perform actions on which it is programmed but ML machines can make their own decisions. Based on the type of Input data available, the ML algorithms are divided into two categories: 1)Supervised Machine Learning 2)Unsupervised Machine Learning

1)Supervised Machine Learning: This type of ML algorithms takes data with labels as input i.e. the data whose class / category is already known. The machine will be trained first using this labelled data so that the machine will be able to predict the class of a new instance whose label is unknown. Classification and Regression are the types of Supervised ML algorithms.

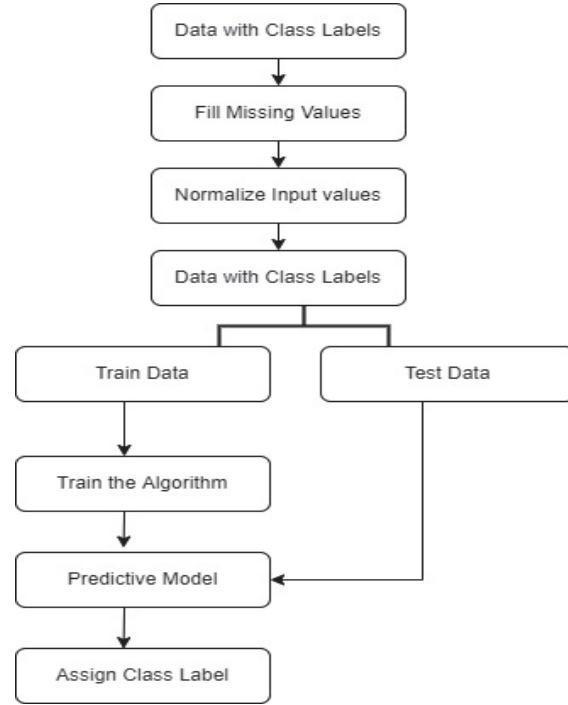


Fig. 1. Steps in ML algorithm

Classification is a technique which takes the new instance as input and it is able to assign the new instance to a particular category. Example of classification are Support Vector Machine, K-Nearest Neighbor classification and Decision Trees. Regression is a type of Supervised ML algorithms which tries to find the relationship between the input variable and the output variable. Using the relationship, regression algorithms finds the output value for a new input value. Example of regression are Linear regression and Logistic regression.

Figure 1 shows the outline of the working a supervised ML algorithm.

2)Unsupervised Machine Learning: This type of ML algorithms takes data without label and analyze those data. This algorithm aims at finding the hidden patterns inside the data. Clustering is an example for unsupervised machine learning algorithms. Clustering takes data with unknown labels as input and tries to find the similarity between data points. K-means clustering and Hierarchical clustering are some examples of clustering.

Artificial Neural Network:

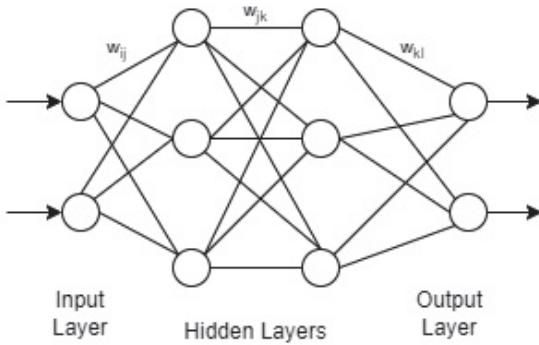


Fig. 2. A simple ANN with 2 hidden layer

Artificial Neural Network (ANN)[22] is a type of supervised ML algorithm whose input will be the collection of labelled data. ANN is basically built on the concept of the human brain which is composed of billions of nerve cells called neurons. These neurons will be interconnected and are used in the processing of records. Similarly, ANN is composed of nodes / neurons which are interconnected to each other. The collection of neurons will be organized into layers: input layer, hidden layer and output layer. Figure 2 shows a simple ANN with two hidden layers.

i.Input Layer: There will be only one input layer and it will have group of neurons. The raw input from the world will be in the format of text, image, signal, voice etc., This raw input is fed into the Input layer.

ii.Hidden Layer: There can be one or more number of hidden layers. All the input along with the corresponding output values will reach the neurons. The neurons will apply a

Summation function as well as activation function to the input values. Depending on the activation function, some of the neurons will fire and send output to next layer. This process continues until the values reaches the output layer.

iii.Output Layer: There will be only one output layer. This layer is responsible for accurate classification of a new instance.

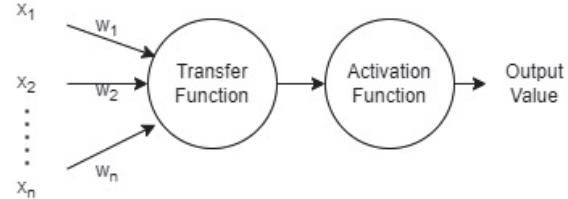


Fig. 3.Working of Single Neuron

Figure 3 shows the working of a single neuron. All input values will be sent to all neurons along with corresponding weights. The neuron will first apply the transfer function to the input values and weights which is given by equation 1,

$$f(X) = \sum_{i=0}^n (x_i * w_i) + bias \quad (1)$$

The result of the transfer function will be given to activation function which decides whether the neuron should send the calculated value to the next layer of neurons. The main function of activation function is to bring non-linearity to the input values and weights. The following are some types of activation functions:

i.Binary step function: Here, the output from the transfer function will be compared to a threshold. If it is greater than or equal the threshold, the neurons will fire. The function is given by the equation 2,

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (2)$$

ii.Sigmoid function: It takes any range of values as input and maps it to a value between 0 and 1. This type of function is used when we need the output as a probability value as equation 3.

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

iii.Tanh function: It takes any range of values as input and maps it to a value between -1 and 1 as given in equation 4.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

iv.ReLu function: It uses the concept of derivation and it can also be used for back propagation. It produces more efficient outputs when compared to other functions using equation 5.

$$f(x) = \max(0, x) \quad (5)$$

IV. ARTIFICIAL NEURAL NETWORK IN THE FIELD OF BIO-FUELS

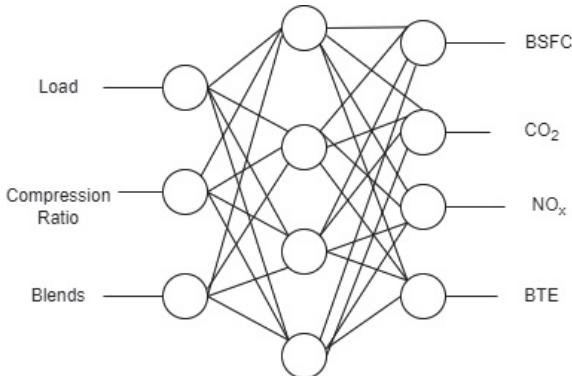


Fig. 4.ANN modelled with input and output parameters

The experiment is carried out with blends of diesel, jatropha oil and neem oil in three different mixture: M0[diesel], M1[50%diesel,40% jatropha, 10% neem], M2[50%diesel,30% jatropha, 20% neem] and M3[50%diesel,20% jatropha, 30% neem]. Transesterification and Esterification process has been carried out for producing the bio-fuels. Esterification is used to increase the volume of bio-fuel produced and through this process the impurities present in the vegetable oil will be removed. In these processes we have produced bio-fuel by making the chemical reaction between ethanol, jatropha oil and neem oil.

Among the input parameters, load of the engine is varied to analyse the output. The different load values considered are 25%, 50%, 75% and 100% and the compression ratios considered are 17.5%, 18.5%, 19.5%. The output parameters considered are Emission of Nitrogen Oxide (NO_x), Emission of Carbon-dioxide (CO₂), Break Specific Fuel Consumption (BSFC) and Brake Thermal Efficiency (BTE).

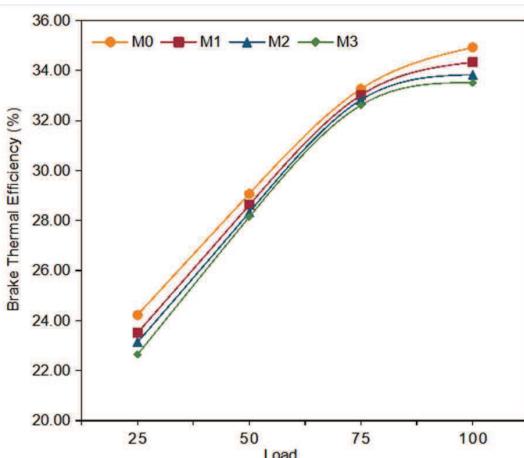


Fig. 5. BTE vs Load

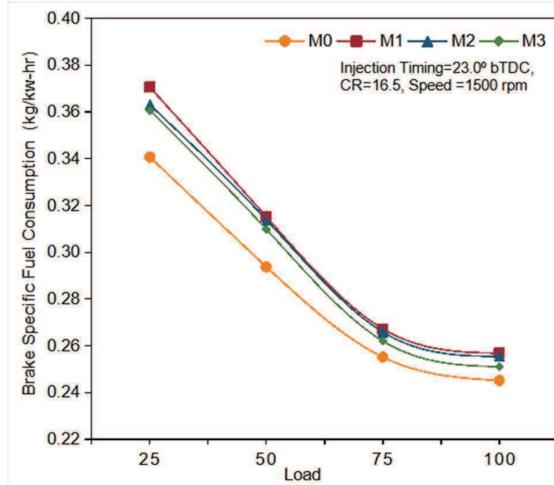


Fig. 6. BSFC vs Load

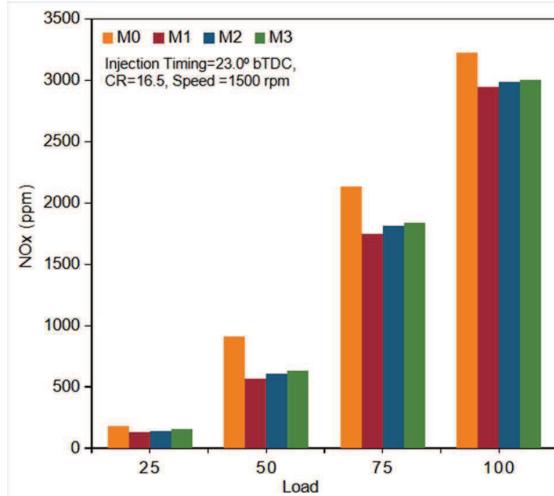


Fig. 7.NO_x Emission vs Load

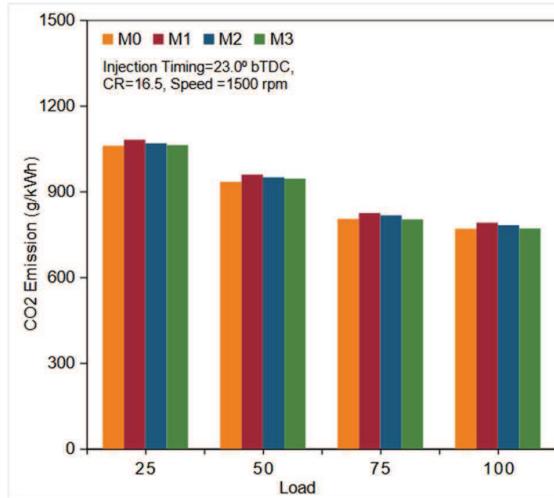


Fig.8 CO₂ Emission vs Load

Figure 4 shows the model of ANN built for evaluating the input parameters and output parameters. The activation function as well as the learning rate of the neural network is varied to evaluate the performance. The weights of

the neural network is adjusted according to the feedback from the output layer. The weight adjustments are done in order to reduce the overall error in the output produced.

Figure 5 shows the Brake Thermal Efficiency (BTE) against the various loads for 4 different blends as input. Figure 6 shows the Brake Specific Fuel Consumption (BSFC) against the various loads for 4 different blends as input. Figure 7 shows the NO_x emission against the various loads for 4 different blends as input. Figure 8 shows the CO₂ Emission against the various loads for 4 different blends as input. From the results obtained, we can analyze the effect of using various blends of jatropha oil and neem oil as an alternate for petroleum based fuels.

V. CONCLUSION

Bio-Diesels are becoming popular since they are renewable and environment friendly. Many researches have been carried out to predict the bio-fuel which is giving good engine performance. In this work, three different mixtures of diesel, jatropha oil and neem oil has been considered for experimentation. Loads are taken as 25%, 50%, 75% and 100% and compression ratios is 16.5%. The learning rate of ANN was also varied and the activation function was also varied. The output parameters considered are Emission of Nitrogen Oxide (NO_x), Emission of Carbon-dioxide (CO₂), Break Specific Fuel Consumption (BSFC) and Brake Thermal Efficiency (BTE). Experiments shows that ANN has given better optimization of engine parameters.

REFERENCES

- [1] Jahirul, M.I. & Brown, Richard & Senadeera, Wiji & O'Hara, Ian & Ristovski, Zoran. (2013). The Use of Artificial Neural Networks for Identifying Sustainable Biodiesel Feedstocks. *Energies*. 6. 3764-3806. 10.3390/en6083764.
- [2] Bhatt, Aditya & Shrivastava, Nitin. (2021). Application of Artificial Neural Network for Internal Combustion Engines: A State of the Art Review. *Archives of Computational Methods in Engineering*. 10.1007/s11831-021-09596-5.
- [3] Agrawal, Tanmaya & Gautam, Raghvendra & Agrawal, Sudeeksha & Singh, Vishal & Kumar, Manish & Kumar, Saket. (2020). Optimization of engine performance parameters and exhaust emissions in compression ignition engine fueled with biodiesel-alcohol blends using taguchi method, multiple regression and artificial neural network. *Sustainable Futures*. 2. 100039. 10.1016/j.sfr.2020.100039.
- [4] Narendranathan, S K, Sudhagar, K, Karthikeyan, R., "Optimization of engine operating parameters suitable for punnai oil application in ci engine using grey relational method. *Energy & Environment*, vol. 30, 732-751, 10.1177/0958305X18813603.
- [5] Sayyed, Siraj & Das, Randip & Kulkarni, Kishor. (2021). Performance assessment of multiple biodiesel blended diesel engine and NO_x modeling using ANN. *Case Studies in Thermal Engineering*. 28. 10.1016/j.csite.2021.101509.
- [6] Kumar, L. & Ramakrishnan, Prakash. (2020). Validation of performance and emissions of a CI engine fueled with calophyllum inophyllum methyl esters using soft computing technique. *Fuel*. 266. 117070. 10.1016/j.fuel.2020.117070.
- [7] Gnanasekaran, Sakhivel. (2016). Prediction of CI engine performance, emission and combustion characteristics using fish oil as a biodiesel at different injection timing using fuzzy logic. *Fuel*. 183. 214-229. 10.1016/j.fuel.2016.06.063.
- [8] Shrivastava, Pankaj & Salam, Satishch & Verma, Tikendra & Samuel, Olusegun. (2019). Experimental and empirical analysis of an IC engine operating with ternary blends of diesel, karanja and roselle biodiesels. *Fuel*. 262. 1-17. 10.1016/j.fuel.2019.116608.
- [9] Hariharan, N. & Senthil, V. & M, Krishnamoorthi & Karthic, S.V.. (2020). Application of artificial neural network and response surface methodology for predicting and optimizing dual-fuel CI engine characteristics using hydrogen and bio fuel with water injection. *Fuel*. 270. 117576. 10.1016/j.fuel.2020.117576.
- [10] Yang, Fubin & Cho, Heejin & Zhang, Hongguang & Zhang, Jian & Wu, Yuting. (2018). Artificial neural network (ANN) based prediction and optimization of an organic Rankine cycle (ORC) for diesel engine waste heat recovery. *Energy Conversion and Management*. 164. 15-26. 10.1016/j.enconman.2018.02.062.
- [11] Ganapathy, Thirunavukkarasu & Krishnan, Murugesan & Gakkhar, Rakesh. (2009). Performance optimization of Jatropha biodiesel engine model using Taguchi approach. *Applied Energy*. 86. 2476-2486. 10.1016/j.apenergy.2009.02.008.
- [12] Kurtgoz, Yusuf & Karagoz, Mustafa & Deniz, Emrah. (2018). Biogas engine performance estimation using ANN. *Engineering Science and Technology, an International Journal*. 20. 10.1016/j.jestch.2017.12.010.
- [13] Channapattana, Shylesha & Pawar, Abhay & Kamble, Prashant. (2017). Optimisation of operating parameters of DI-CI engine fueled with second generation Bio-fuel and development of ANN based prediction model. *Applied Energy*. 187. 84-95. 10.1016/j.apenergy.2016.11.030.
- [14] Deb, Madhujit & Majumder, Pinki & Majumder, Arindam & Roy, Sumit & Banerjee, Rahul. (2016). Application of artificial intelligence (AI) in characterization of the performance-emission profile of a single cylinder CI engine operating with hydrogen in dual fuel mode: An ANN approach with fuzzy-logic based topology optimization. *International Journal of Hydrogen Energy*. 41. 10.1016/j.ijhydene.2016.07.016.
- [15] Bhowmik, Subrata & Paul, Abhishek & Panua, Rajsekhar & Ghosh, Subrata & DEBROY, DURBADAL. (2018). Performance-Exhaust Emission Prediction of Diesenosol Fueled Diesel Engine: An ANN Coupled MORSIM Based Optimization. *Energy*. 153. 10.1016/j.energy.2018.04.053.
- [16] Gnanasekaran, Sakhivel & Sivaraja, C.M. & Ikua, Bernard. (2018). Prediction OF CI engine performance, emission and combustion parameters using fish oil as a biodiesel by fuzzy-GA. *Energy*. 166. 10.1016/j.energy.2018.10.023.
- [17] Dey, Suman & Reang, Narath & Majumder, Arindam & Deb, Madhujit & Das, Pankaj. (2020). A hybrid ANN-Fuzzy approach for optimization of engine operating parameters of a CI engine fueled with diesel-palm biodiesel-ethanol blend. *Energy*. 202. 117813. 10.1016/j.energy.2020.117813.
- [18] Syed, Javed & Yedithasatayam, Murthy & Baig, Rahmath & Rao, D.. (2015). Development of ANN model for prediction of performance and emission characteristics of hydrogen dual fueled diesel engine with Jatropha Methyl Ester biodiesel blends. *Journal of Natural Gas Science and Engineering*. 26. 549-557. 10.1016/j.jngse.2015.06.041.
- [19] Dehghani, Masoud & Ghobadian, Barat & Najafi, Gholamhassan & Sabzimaleki, Mohammmdreza & Jaliliatabar, Farzad. (2015). Performance and Exhaust Emissions of a SI Two-stroke Engine with Biolubricants Using Artificial Neural Network.
- [20] Hoang, Anh & Nizetic, Sandro & Ong, Hwai Chyuan & Tarelko, Wieslaw & Viet, Pham & Tri Le, Hieu & Chau, Minh & Nguyen, Xuan Phuong. (2021). A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels. *Sustainable Energy Technologies and Assessments*. 10.1016/j.seta.2021.101416.
- [21] Patil, Rajendra. (2012). ANN & It's Application To IC Engine.
- [22] Liu, Zhiqiang & Zuo, Qingsong & Wu, Gang & Li, Yuelin. (2018). An artificial neural network developed for predicting of performance and emissions of a spark ignition engine fueled with butanol-gasoline blends. *Advances in Mechanical Engineering*. 10. 168781401774843. 10.1177/168781401774843.

Design of an All-Digital Phase-locked loop in a 130nm CMOS Process using open-source tools

Charaan S

Department of Electronics Engineering
MIT Campus, Anna University
Chennai, India
Email: kumarcharaan27@gmail.com

Nalinkumar S

Department of Electronics Engineering
MIT Campus, Anna University
Chennai, India
Email: nalintes36@gmail.com

Elavarasan P

Department of Electronics Engineering
MIT Campus, Anna University
Chennai, India
Email: elavarasanp002@gmail.com

Prakash P

Department of Electronics Engineering
MIT Campus, Anna University
Chennai, India
Email: prakashp79@gmail.com

Kasthuri P

Department of Electronics Engineering
MIT Campus, Anna University
Chennai, India
Email: kasthurip94@gmail.com

Abstract—In this paper, an implementation of an ‘All-Digital Phase-Locked Loop’ has been elucidated, which includes the RTL level synthesis of the design, and the GDS level synthesis, obtained from the RTL to GDSII flow. The implemented design consists of a Phase Frequency Detector, a K-Counter, a Digital Controlled Oscillator, with a Divide by N Counter in a feedback loop. These individual components were realized using Verilog Hardware Description Language, and an appropriate testbench was written to view the waveforms and check for their correctness using GTKWave. Later, the RTL Level Design of this implementation was obtained, and, the RTL to GDSII Flow was initialized. The tool used for executing this process is OpenLane, which is an open-sourced, automated RTL to GDSII flow based on several components like OpenRoad, Yosys, Magic, Klayout, etc. The different stages of the RTL to GDSII Flow such as Floorplanning, Placement, CTS, and Routing were performed, and the GDS and LEF files were obtained successfully. The Gate Level Simulation was done for the post-layout stage, by obtaining the synthesized design from the layout stage. The obtained GDSII file is incorporated with the open-source Sky130 Process Design Kit and can be sent to foundries that fabricate 130 nm CMOS Processes.

Keywords—ADPLL, RTL Design, RTL to GDSII Flow, OpenLANE, Sky130 PDK, 130 nm

I. INTRODUCTION

The idea of synchronizing two clock signals came into existence as early as the 17th century, when people were involved in developing mechanisms for synchronizing pendulum clocks and tuning forks. Later, with the help of the innovations going on with oscillator technology, a standard synchronization circuit was developed, in which the phase of two signals was synchronized. This circuit was widely known as the Phase-Locked Loop (PLL), and it has seen a lot of innovations in recent years. This circuit is widely implemented in data transmission systems, where the transmitter and receiver are required to be synchronized in order to access the transmitted data. Here, the PLL generates a signal from a local oscillator that is spectrally pure, and this enables the translation

of frequency between the transmitter and the receiver [1]. Analog PLLs are mostly implemented for these applications, but they come with limitations such as noise coupling and signal shading [2]. Also, the PLL requires scaling of the reference voltage for every new CMOS process technology [2].

In order to address the shortcomings mentioned above, a purely digital version of a PLL, more commonly known as an All-Digital PLL is being implemented in various applications. One great advantage that ADPLLs offer is its non-requirement to scale the circuit for every new CMOS process technology. Also, there is no need for a Mixed-Mode Silicon Process [2]. The portability over different CMOS processes becomes feasible by realizing the circuit using an appropriate Hardware Description Language (HDL). This enables the design to become applicable to any standard cell library [3]. A typical analog PLL comprises a phase detector, loop filter, and a Voltage-Controlled Oscillator (VCO). In ADPLLs, the phase detectors and analog filters are replaced with time-to-digital converters and digital loop filters, thereby significantly reducing area and providing higher compatibility with other digital blocks [4]. ADPLLs are synthesized primarily using digital design flows. The increasing complexity of digital circuits has led to a corresponding increase in the sophistication of the associated stages of digital design such as synthesis, layout, and verification, which are now being performed using Electronic Design Automation (EDA) tools [4]. On the other hand, for analog PLLs, a fully-custom layout is required as the performance is highly dependent on the layout of that particular analog design. Hence, it becomes advantageous in a commercial perspective to have the analog components implemented in digital form, and then the designing for the same is done using CAD tools [4].

This paper presents, the design, and layout of an ADPLL, primarily done using open-source tools. The two main advantages of designing hardware using open-source tools are reduction of design time and reduction of design cost [5]. Also, hardware designed using open-source tools can be published in

open-source hardware libraries that come with community support [5], which gives rise to improvements in the designs. It also enables fabrication of the designed hardware at a reduced cost using open-source Process Design Kits. One such PDK is the SkyWater PDK [6] which is a collaboration between Google and SkyWater Technology Foundry. The technology node used here is the Sky130 process node, which is a 130nm CMOS technology node [6].

The remainder of this paper is organized as follows. The second section of the paper describes the architecture of the designed ADPLL and the individual components that make up the design. It also presents the pre-layout simulation of the same. Section III gives a brief overview of OpenLANE [5] and describes the RTL to GDSII Flow [7]. This corresponds to the layout stage of the ADPLL design, and the GDSII file obtained at the end of this process can be sent to foundries for fabrication.

II. RELATED WORK

An All-Digital Phase-Locked loop is a circuit whose elements are purely digital, unlike their analog counterparts. They are realized using Hardware Description Languages (HDL), thereby making them fully customizable and compatible with the required application. They eliminate the limitations in the supply voltage of analog PLLs and are also fully synthesizable, thus enabling them to be implemented in any process node across all foundries. This results in a decreased time to market, hence ADPLLs are nowadays widely used, and are slowly replacing analog PLLs.

One way of designing ADPLLs is using the Rapid Single Flux Quantum (RSFQ) technology, as proposed by Cong and Pedram [8]. It is an emerging technology that employs superconducting devices for processing the digital signals that are essential for the circuit. This technology is proved to be compatible with the circuits that are designed using the conventional CMOS technology, and circuits that are designed using SFQ technology come with low power consumption and high clock frequencies. Here, the Digital Controlled Oscillator (DCO) is designed using a tunable delay cell, and the circuit consumes a total of 0.45 mW and occupies an area of 7.76 mm².

Cheong and Kim [9] have presented an ADPLL circuit using Triple-Stage Phase-Shifting (TSPS), thereby enabling fast-locking. This is achieved by first implementing a Binary Search Algorithm (BSA), followed by a TSPS scheme that supports the reduction of phase offset between the feedback and reference clocks and at the same time eliminates frequency peaking. This circuit was designed in a 28nm CMOS process and the amount of power dissipated is given as 2.89 mW.

Deng et al. [10], have implemented a fast-locking ADPLL using the Bisection Method. Here, the DCO possesses a wide tuning range and a higher resolution. This design comes with a control circuit that determines the frequency control word required for the DCO. This design is capable of reducing the locking time by around 80%, and the core of the circuit occupies 0.04 mm². The circuit comes with a power consumption of 29.48 mW, and also has a wide tuning range.

The above-mentioned papers have carried out their research using commercial tools and proprietary software. In this paper, the design and analysis of an ADPLL circuit have been exclusively done using open-source tools. The primary advantage of using open-source tools lies in the fact that the source code of the design made using such tools, can be easily licensed as Intellectual Property (IP), and the same can be re-used or integrated as components in other larger systems. The obtained IP specifications can then be shared with sites that facilitate Free and Open-Source collaboration.

Using various open-source EDA tools, the design of an ADPLL was made in a 130nm CMOS Process, which comes with an area of 0.58 mm² and power dissipation of 0.201 mW.

III. THE ADPLL ARCHITECTURE

A typical Phase-Locked Loop comes with a Phase Detector, Low Pass Filter, Voltage Controlled Oscillator (VCO), with a Divide by N Counter in a feedback loop between the Phase Detector and VCO. In an ADPLL, the analog components are replaced with digital circuits, thereby providing better portability [11] and scalability. The signals that are transmitted between each block are purely digital and could be binary or word signals such as the output coming from a counter or data register [11]. It consists of an All-Digital Phase Detector, K - Counter Loop Filter and a Digitally Controlled Oscillator (DCO) that comes with an increment/decrement counter and a divide by N counter [12]. The block diagram of the architecture of an ADPLL is shown in Figure 1. Each of the individual components is explained in the following sub-sections.

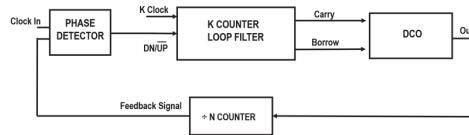


Fig. 1. Block Diagram of ADPLL

A. Digital Phase Detector

Also known as a phase comparator, the digital phase detector compares two digital inputs which are in the form of square waves, and generates an error signal based on the phase of the two inputs. This error signal is the difference in phases between the two input signals, and in an ADPLL, the two input signals passed to the Digital Phase Detector are the reference signal and the output coming from the DCO. A simple way to produce a signal that carries the phase difference between two input signals is to pass those two inputs to an XOR gate [13] as shown in Figure 2. The XOR gate, by virtue, will be low if two signals are perfectly in-phase with each other, and will be high if there exists a phase difference. The duration of the high and low levels of the XOR gate depends on the difference in phase occurring at each instant of the period. The Digital Phase Detector hence compares the output coming from the DCO with

the reference signal and produces an error signal which is then sent to the K-Counter Loop-Filter.

The Phase Detector Gain (K_{PD}) for the XOR based Digital Phase Detector is given as

$$K_{PD} = \frac{V_{DD}}{\pi} \quad (1)$$

Where V_{DD} refers to the logic level HIGH of the XOR Phase Detector. The logic level LOW is taken as zero [11].

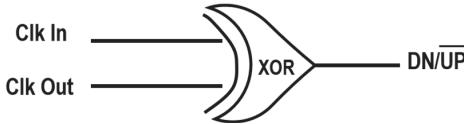


Fig. 2. XOR based Digital Phase Detector

B. K-Counter Loop-Filter

In conventional PLLs, an analog loop filter is implemented in order to filter out the higher-order frequencies coming out of the Phase Detector. This loop filter is usually a low-pass filter and is implemented using a resistor-capacitor network [11]. In an ADPLL, the resistor-capacitor networks are replaced with up and down counters as shown in Figure 3, which have a range from 0 to $K-1$, where K is the modulus value of the counter, which indicates the number of states that the counter will be able to count. The configuration of the counter at any particular time is set by the DN/UP signal. If the DN/UP signal is at logic-level HIGH, then the down-counter is enabled, and if the DN/UP signal is at logic-level LOW, then the up-counter is enabled. Here, the value of K is taken as 256 [14], hence the counter ideally starts counting from 0 to 255 or 00 to FF in hexadecimal, if the counter is configured to count upwards. Similarly, the down counter counts from FF to 00. When the count at any configuration, exceeds $K-1$, both the counters are reset to 0. The K-Counter gives two outputs - carry and borrow. At any instant, if the value of the down counter exceeds or becomes equal to $K/2$, the borrow gives out '1'. Likewise, the carry gives out '1' if the value of the up-counter becomes greater or equal to $K/2$. The carry and borrow signals are given by the MSBs of up and down counter respectively. K-Counter plays a crucial role in operating the DCO, and the positive edges of the carry and borrow signals are used to control the frequency of the DCO.

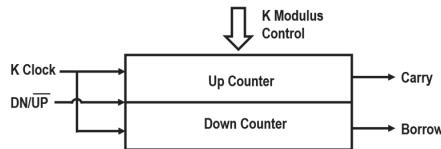


Fig. 3. K-Counter Loop-Filter

C. Digital Controlled Oscillator (DCO)

In analog PLLs, Voltage Controlled Oscillators are employed to generate the output voltage signal of the PLL, whose

frequency is varied in accordance with a control voltage V_c . The operating frequency is given by the expression $f = KV_c$, where K is the Sensitivity of the VCO given in rads/V and $V_c > 0$. This function of generating the output signal is governed by the DCO in an ADPLL. The overall design of a DCO consists of an Increment/Decrement Counter and a Divide by N counter. Here the DCO modifies the frequency of the output signal based on the carry and borrow signals, at that particular instant. The carry and borrow signals from the K-Counter Loop-Filter are sent to the DECR and INCR ports of the ID Counter respectively. The output (OUT) of the ID Counter is now used as the clock for the Divide by N counter. If at any instant, the carry and borrow signals are at logic level LOW, then the ID Counter divides its output (OUT) by 2 on the positive edge of the clock sent to the ID Counter. The Divide by N counter scales down the frequency of the reference signal generated from a high-frequency oscillator. The frequency of the output signal of DCO is given by the formula

$$f_{DCO} = \frac{f_{REF}}{N} \quad (2)$$

Where f_{DCO} is the frequency of the output signal of DCO and f_{REF} is the frequency of the reference signal given to the Phase Detector [11]. The Divide by N Counter hence controls the frequency that needs to be fed to the Phase Detector. The DCO block is shown in Figure 4.

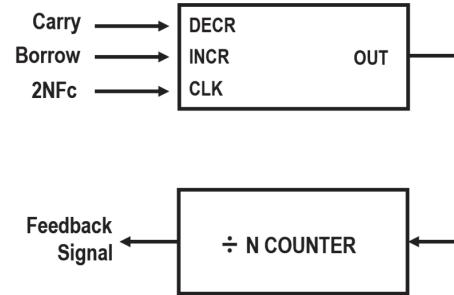


Fig. 4. Digital Controlled Oscillator

D. Pre - Layout Simulation

The RTL design of the components mentioned above is realized using Verilog HDL, and the simulation for the same is done using GTKWave, which is an open-source waveform viewer. This simulation was done, prior to the RTL to GDSII Flow, and is shown in Figure 5. Here, it is observed that for each level of clk_{out} , the clk_{out_8x} signal occupies exactly 8 cycles throughout the entire duration of that level. Also, there is a clear difference in phase between the clk_{in} and clk_{out} signals.



Fig. 5. Pre-Layout simulation results

IV. LAYOUT OF ADPLL

A. Introduction to OpenLANE

OpenLANE is an open-source, automated RTL to GDSII Design Flow developed by The OpenROAD Project. It aims to reduce the shortcomings faced by designers in terms of cost and expertise while allowing the implementation of hardware designs using advanced technologies [15]. It is potentially a start of delivering tapeout-capable hardware designs along with Google SkyWater PDK, thereby enabling hardware designers to fabricate their designs without relying on commercial EDA tools and technology nodes. The entire flow of OpenLANE, consists of several stages, wherein each step is made to run sequentially. The different stages of OpenLANE include synthesis, floorplanning Placement, Clock Tree Synthesis (CTS), Routing, GDSII generation, DRC, and LVS checks [7]. OpenLANE comes with several other opensource components, that are involved in executing each of the steps associated with the RTL to GDSII Flow, such as Yosys, Netgen, Magic, and Klayout.

B. The process

The RTL to GDSII Flow is initiated by bringing the RTL design of the digital circuit and loading it onto OpenLANE. The software later gives out the appropriate files for each stage of the design flow. This whole process is automated; hence it does not require manual intervention. Each stage is executed sequentially, and the process terminates with the generation of the GDSII file. Along with the generation of the GDSII file, the results of the Design Rule Checking (DRC) and Layout Versus Schematic (LVS) are also presented. The DRC is performed using Magic and the LVS check is performed using Netgen. The obtained GDSII file format is the widely accepted standard for Integrated Circuit layout and fabrication and can be viewed either using Magic or Klayout.

For the RTL Design of the ADPLL, the appropriate GDSII file was generated successfully as a result of the automated process of OpenLANE and its underlying components. Figure 6 shows the DRC result viewed through Magic, while Figure 7 shows the GDSII file obtained through Klayout.

V. POST - LAYOUT

A Gate-Level Simulation (GLS) is highly essential, in order to verify the correctness and functionality of the design after layout and synthesis. Figure 8 shows the GLS of the ADPLL

design obtained after the RTL to GDSII flow. The `clk_out_8x` signal contains 8 cycles for the duration of each level of `clk_out`, and thus appears to have the same functionality as that of the pre-layout design. The specifications of the designed ADPLL circuit are summarized in Table I.

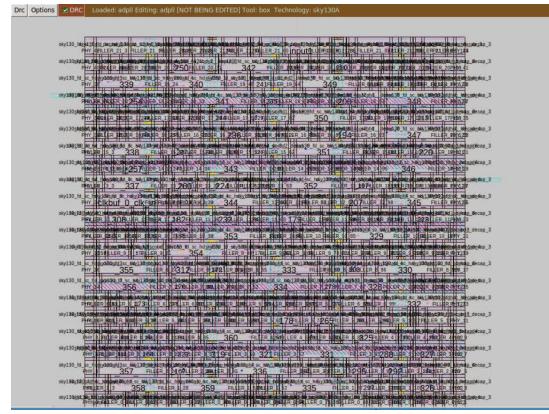


Fig. 6. DRC result of the RTL to GDSII Flow. It is observed that there are no DRC violations.

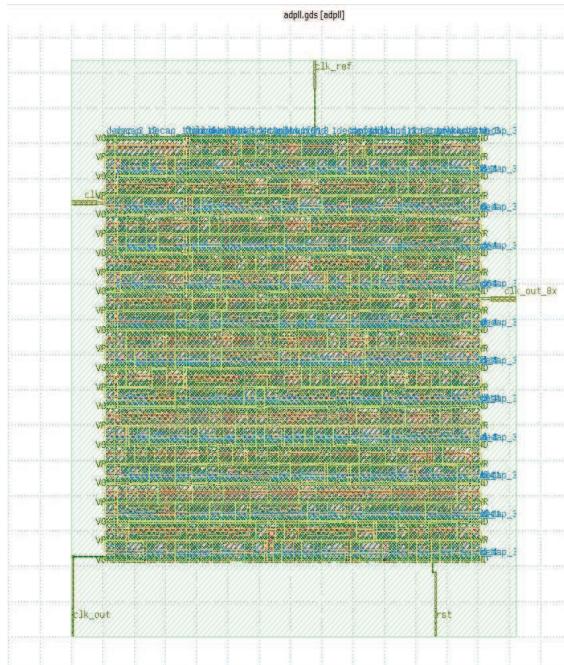


Fig. 7. GDSII file viewed through Klayout

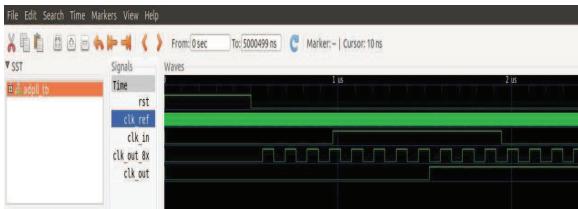


Fig. 8. Gate-Level Simulation of the designed ADPLL

TABLE I
SPECIFICATIONS FOR THE DESIGNED ADPLL

Parameters	
Technology (nm)	130 nm CMOS
Area (mm ²)	0.58
Power Consumption	0.201mW
Cycle to Cycle jitter	50ps

VI. CONCLUSION

In this paper, an All-Digital Phase-Locked Loop is designed using open-source tools. The ADPLL circuit is modelled using the 130 nm CMOS technology node, taken from Sky130 PDK. First, the individual blocks of an ADPLL are realized using Verilog HDL. Next, the RTL design is simulated and the obtained waveforms are verified. The layout phase is then initiated by launching OpenLANE, and loading the RTL design of the ADPLL onto the flow. The RTL to GDSII Flow is executed in a sequential manner, covering each step such as synthesis, floor planning, placement, CTS, and Routing, and generates the GDSII file, that can be sent to foundries for fabrication purposes. The DRC and LVS check results are analyzed, and later the Gate-Level Simulation (GLS) is done on the ADPLL design, obtained after layout. The parameters of the designed circuit are computed, and the area and power dissipation of the circuit is found to be 0.58 mm² and 0.201 mW respectively. From the obtained values, we can infer that the circuit is approximately 45% more efficient than the other circuits discussed.

ACKNOWLEDGMENT

We would like to extend our profound thanks to Dr. M. Ganesh Madhan, Head of the Department, Department of Electronics Engineering, Madras Institute of Technology, Anna University, for his continued support and guidance and for providing us with the facilities required to work on this project.

REFERENCES

- [1] S. Aniruddhan, S. Shekhar and D. J. Allstot, "A CMOS 1.6 GHz DualLoop PLL With Fourth-Harmonic Mixing," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 5, pp. 860-867, 2011.
- [2] S. Moorthi, D. Meganathan, D. Janarthanan, P. Praveen Kumar and J. Raja Paul Perinbam, "Low jitter all digital phase locked loop based clock generator for high speed system on-chip applications," *International Journal of Electronics*, 2009, 96:11, pp. 1183-1189.
- [3] C. Wang, C. Huang, and S. Tseng, "A low-power ADPLL using feedback DCO quarterly disabled in time domain," *Microelectron. J.* 39, 5 (May, 2008), pp 832–840.
- [4] Y. Park and D. D. Wentzloff, "An all-digital PLL synthesized from a digital standard cell library in 65nm CMOS," *IEEE Custom Integrated Circuits Conference (CICC)*, 2011, pp. 1-4.
- [5] M. Shalan and T. Edwards, "Building OpenLANE: a 130nm openroadbased tapeout-proven flow", In *Proceedings of the 39th International Conference on Computer-Aided Design (ICCAD '20)*. Association for Computing Machinery, New York, NY, USA, Article 110, pp 1-6.
- [6] T. Ansell and M. Saligane, "The Missing Pieces of Open Design Enablement: A Recent History of Google Efforts : Invited Paper," *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, 2020, pp. 1-8.
- [7] J. Lu and B. Taskin, "From RTL to GDSII: An ASIC design course development using Synopsys® University Program," *2011 IEEE International Conference on Microelectronic Systems Education*, 2011, pp. 72-75.
- [8] H. Cong and M. Pedram, "All-Digital Phase-Locked Loop in Single Flux Quantum Circuit Technology," in *IEEE Transactions on Applied Superconductivity*, vol. 32, no. 3, pp. 1-8, April 2022.
- [9] H. H. Cheong and S. Kim, "A Fast-Locking All-Digital PLL With Triple-Stage Phase-Shifting," in *IEEE Access*, vol. 9, pp. 160224160237, 2021.
- [10] Deng, Xiaoying, Li, Huazhang, Zhu and Mingcheng, "A Novel FastLocking ADPLL Based on Bisection Method," in *Electronics*, 2021.
- [11] A. Gothandaraman, "Design and Implementation of an All Digital Phase Locked Loop using a Pulse Output Direct Digital Frequency Synthesizer," Master's Thesis, University of Tennessee, 2004.
- [12] Texas Instruments, "Digital Phase-Locked Loop Design Using SN54/74LS297," *SN54/74LS297 Datasheet*, March 1997.
- [13] A. Singhal, C. Madhu and V. Kumar, "Designs of All Digital Phase Locked Loop," *Recent Advances in Engineering and Computational Sciences (RAECS)*, 2014, pp. 1-5.
- [14] Rajni Singla, Rekha Yadav, "Design an All Digital PLL with Ripple Reduction Technique," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN: 2278-3075, vol. 8, issue-10, 2019.
- [15] T. Ajayi et al., "INVITED: Toward an Open-Source Digital Flow: First Learnings from the OpenROAD Project," *56th ACM/IEEE Design Automation Conference (DAC)*, 2019, pp. 1-4.

A Term Weight Measure based Approach for Author Profiling

Karunakar Kavuri

Research Scholar, Department of CSE

VelTech Rangarajan Dr.Sagunthala R&D Institute

of Science and Technology

Avadi, Chennai, India

karunakar.mtech@gmail.com

M Kavitha

Professor, Department of CSE

VelTech Rangarajan Dr.Sagunthala R&D Institute

of Science and Technology

Avadi, Chennai, India

kavitha@veltech.edu.in

Abstract— Author Profiling (AP) is a task of determining the demographic features like Age, Gender, Nativity language, location, Personality traits etc. about the author of a document. AP is used in different applications such as marketing linguistic profile, security and forensic science. The researchers proposed different types of solutions to author profiling based on stylistic features, content based features and deep learning techniques. The content based features proved their significance to improve the performance of author profiles prediction. Several approaches faced a problem of high dimensionality of features when experimented with content based features. The researchers used feature selection algorithms to identify the important features for experimentation. In this work, a Term Weight Measure (TWM) based approach is proposed for author profiling problem. In this approach, the important features are identified by using Feature Selection (FS) algorithm. After features are identified, the next important task is representation of document with identified features. The documents are represented as vectors and computation of each feature value in the vector representation is another important research task. TWMs are used to determine the importance of a feature in the vector representation. In the proposed approach, we proposed a new TWM based on the way the terms are distributed in corpus of documents. The proposed term weight measure performance is compared with different existing TWMs. The PAN competition 2014 reviews dataset is used for age and gender prediction of the author. Two Machine Learning (ML) algorithms such as Random Forest (RF) and Support Vector Machine (SVM) are used to evaluate the proposed term weight measure based approach. The experimental results attained in this work for age and gender prediction are good when compared with several popular solutions to author profiling.

Keywords— Author Profiling, Age Prediction, Gender Prediction, Feature Selection Algorithm, Term Weight Measure, Machine Learning Algorithms

I. INTRODUCTION

The information in the form of text is exponentially enhancing in the internet through different types of websites like blogs, forums, reviews and other social media websites. The crimes also enhancing with the increment of textual information like harassing messages, threatening mails and fake profiles. Knowing the author details about textual

information becomes an important task to reduce the crimes in the internet. Authorship Analysis is one research area to extract the information of the author from their written texts. Authorship analysis was classified into three classes such as Authorship Verification (AV), Authorship Attribution (AA) and Authorship Profiling [AP] [1].

The AV method verifies whether the new text is written by a suspected author or not by examining different texts of suspected author. AA determines the author of a new text by examining multiple authors' text [2]. The Authorship Profiling task (AP) is to extract demographic aspects like gender, age, location, occupation, socio-economic level and native language, level of well-being, personality traits or educational background of a person from their texts [3]. At the beginning of the AP task, formal texts such as newspapers, books or magazines were analyzed to determine the aspects of their authors. In recent years, researchers concentrated on determining the profile of people through their social network accounts. Authorship Profiling is used in various real-time applications like marketing, forensics, security etc.

The use of author profiling was changed from only be used in the forensic investigation and internet security, also to be applied in targeted marketing, advertisement and educational domain [3]. Companies interested in obtaining the knowledge of what describe the people that like or dislike their product. In marketing domain, the expert's interest is to know the identity and demographic characteristics of the various users, with the intention of directing the advertising for exploiting the product in a better way. Authorship Profiling helped in security to narrow the potential number of suspects, or even help rule out potential suspects. In social media websites, most of the people are hiding their original details in their profiles and perform illegal or deceptive acts such as sexual harassment and extortion through textual messages. AP techniques are helpful to determine the basic information of perpetrator. The Forensic linguists make use of linguistic knowledge and authorship profiling techniques to study texts and determining the type of bad behaviors. In educational domain, AP techniques are helpful to determine the exceptional students by examining their styles of writings.

Most of the research works exploited various types of Content based Features (CFs) to distinguish the authors' style of writing. CFs are based on the content bearing terms used by the authors in their text. The corpus contains huge number of content bearing terms, but few terms are more informative terms which are helpful for discriminating the authors. Most of the researchers used feature selection algorithms to determine the best informative features. These

identified features are used for vector representation of documents and these vectors are forwarded to Machine Learning (ML) algorithms to generate the classification model. This model is used to identify the gender and age class label of a new document. In this work, a TWM based approach is proposed for AP. In this approach, the experiment started with FS Algorithm to identify the most relevant features. Once features are extracted, the next step is representing the document with identified features as vectors. The ML algorithms understand only document representation only. In the vector representation of features, the value of a feature influences the performance improvement of proposed approaches. TWMs are used for this purpose to compute the value of a feature in document vector representation. In this proposed approach, a new TWM is proposed by considering the term distributions in different classes of documents in dataset. The gender and age profiles are considered in this work for prediction. The ML algorithms such as RF classifier and SVM are used for generating classification model. This model predicts the accuracy of age and gender prediction. In this work, PAN competition 2014 Reviews dataset is used for age and gender prediction.

This paper is planned in 7 sections. The survey on AP techniques is explained in section 2. The dataset properties are presented in section 3. The section 4 discusses the proposed approach for AP and explained the different TWMs and FSA. The experimental results of proposed approach are analysed in section 5. The section 6 discusses the results of gender and prediction. The section 7 concludes this work with future plans.

II. REVIEW ON EXISTING WORKS OF AUTHOR PROFILING

Author profiling is a technique of predicting traits of one or many authors automatically from their text. Author profiling groups documents of authors based on their content, semantic tags used, topics discussed and similarity of documents. Author Profiling also group the authors based on their writing styles in social media circle. Automatic detection of author profiles from the text has various applications in harassment cases detection, forensic analysis and marketing. The benchmark dataset is required to analyze and developing suitable solutions for author profiling. In recent years, different types of standard measures were generated for evaluating different genres such as blogs, social media, tweets and hotel reviews [4].

The exponential growth of textual content in social media creates highly undesirable problems like propagation of offensive and abusive language in the internet. The author profiling approaches are used in detection of hate speech messages. The existing research suggests that the hateful messages which are propagated by the users, form communities around them and share a set of common stereotypes. The existing popular approaches for detection of hate speech messages majorly depend on semantic and lexical cues of text. Pushkar Mishra et al., proposed [5] a novel approach for hate speech detection which includes the profiling features of Twitter users based on community. They experimented on a dataset that contains 16000 tweets and observed that the proposed approach performance is significantly higher than existing popular works in hate speech detection.

Chiyou Zhang et al., Proposed [6] models for identifying the gender, language variety and age from social media text in the shared task of AP and deception detection in Arabic language. The model is developed by using pre-trained BERT (Bidirectional Encoders Representation from Transformers). They obtained accuracies of 81.67%, 54.72% and 93.75% for gender, age and language variety prediction respectively. Zhang et al., developed [7] different models for predicting age, gender and language variety in the Arabic AP and deception detection shared task. The dataset contains Tweets in Arabic language that was splitted into a training data of 225,000 tweets and a test data of 720,000 tweets. They used multi-lingual BERT-based model with 768 hidden units, 12 layers, 12 attention heads and 110,000,000 parameters. The proposed BERT-based model attained accuracies of 81.67% and 54.72% for gender and age prediction respectively.

In recent past years, the bots which are accounts in social media that were operated automatically, gained substantial importance in worldwide. Some of them are used bots for malicious activities like spreading of disinformation or swaying political elections. Detection of social bots becomes an emerging research area in recent times. Inna Vogel et al., proposed [8] a system by using character n-grams, word unigrams and word bigrams as features. Linear SVM was used to train the system. The proposed model obtained an overall accuracy of 0.91, 0.92 for bot detection in Spanish and English datasets respectively. The model achieved 0.78 and 0.82 accuracies for gender prediction in Spanish and English datasets respectively.

Kowsari et al., experimented [9] on a dataset of Twitter messages in English language for gender prediction of an author. The training dataset of Twitter contains messages written by 1800 females and 1800 males. The test dataset contains messages written by 1200 females and 1200 males. They applied different models like CNN and Random Multi-model DL and using TFIDF and Glove features and the final decision was taken by using a scheme of majority vote. The proposed models attained an accuracy of 0.8633 and F1-score of 0.8583 for gender classification.

Author profiling techniques also used to detect the gender information from images also. Moniek Nieuwenhuis et al., developed [10] a system to participate in AP shared task of PAN 2018. The task is predicting the gender of an author from text and images in Arabic, English and Spanish datasets. They participated in all subtasks. The final system submitted in the competition used Logistic Regression as classifier, word and character n-grams as textual features and proportion, presence and number of faces to detect face emotions as well as selfies as image-based features. The experiment also conducted with word embeddings and observed that the performance of a system affected negatively. The experimental results show that the performance was improved slightly for Spanish and Arabic datasets when text-based features are added to image-based features. They obtained highest accuracies of 80.3% for Spanish dataset using only text based features, 78.7% for Arabic dataset using both image and text based features, 81.2% for English dataset using only text based features on test dataset of PAN 2018.

III. DATASET CHARACTERISTICS

In this work, the experiment conducted on the reviews dataset which is collected from PAN 2014 competition reviews dataset [11]. The dataset contains 4160 documents and two profiles such as gender and age about an author. The gender profile contains two classes such as male and female. Both male and female classes contain 2080 documents each. The age profile contains five classes such as 18-24, 25-34, 35-49, 50-64 and 65+. The characteristics pertaining to dataset is displayed in Table 1. The dataset is balanced in case of gender profile and unbalanced in case of age profile.

TABLE 1. THE REVIEWS DATASET CHARACTERISTICS

Classes / Profiles		Number of Reviews
Gender	<i>Male</i>	2080
	<i>Female</i>	2080
Age	<i>18-24</i>	360
	<i>25-34</i>	1000
	<i>35-49</i>	1000
	<i>50-64</i>	1000
	<i>65+</i>	800

In this work, two ML algorithms such as SVM and RF are used for evaluating the proposed approach. Random Forest (RF) Classifier is a classifier that utilizes both bagging of features and random selection which intern uses ensemble learning technique. Random selection is a process of developing a set of decision trees. From the training data, a decision tree is constructed with replacement is called bagging. Each decision tree acts as base classifier in finding the information about the class label of a new document. Later, a voting process conducted which is part of ensembling. In SVM classifiers, identify the support vectors among the set of instances in the dataset. These support vectors are used to predict the author profiles of unknown document.

The outcomes of machine learning algorithms are represented by using different performance measures such as recall, precision, F1-Score and accuracy for evaluation. In this work, the proposed approach results are displayed in terms of accuracy measure. Accuracy is number of author documents are correctly predicted their gender and age from a set of author documents considered for experimentation.

IV. TERM WEIGHT MEASURES BASED APPROACH FOR AUTHOR PROFILING

In this work, a TWM based approach is proposed for age and gender prediction. The procedure of proposed approach is shown in Fig. 1. In this approach, first, the pre-processing techniques such as stopwords elimination and stemming are applied on the training dataset to remove unwanted data from the dataset. The stopwords are words like articles, prepositions, determiners, and pronouns etc., which are not having any distinguishing power to discriminate the authors' style of writing. The stemming is performed to transform the words into their root form to decrease the unique words count in the dataset. After cleaning the dataset, the next important step is identification of features from the dataset.

A feature is a property of a document which is used to differentiate the given document. Majority of the entities and documents have many features. FSAs are used to

identify the relevant features and for eliminating the redundant features. In this approach, RDC FSA is used to find the important features.

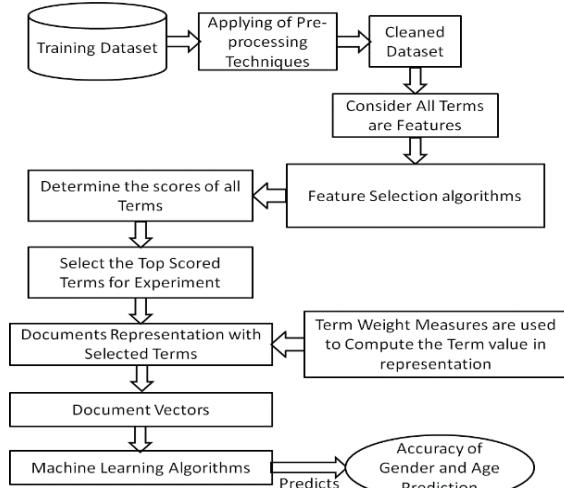


Fig. 1. The Proposed TWMs based Approach

After computing the scores of features by the feature selection algorithms, the top scored features are considered for experiment. In this work, the content based features of terms that are used by the authors in their writings are considered as features. The documents are represented as vectors by using the top ranked terms which are identified in the previous step. In the vector representation, the term value is computed by using TWM. In this work, a new TWM is proposed for computing the term weight. The document vectors are given to MLAs for training. The trained model is used to predict the accuracy of gender and age prediction. The accuracy of gender and age prediction primarily depends on the TWM and FSA that are used in the experimentation.

A. Relative Discriminative Criterion (RDC)

The RDC measure [12] is represented in Equation (1).

$$RDC(T_i, C_j) = \left(\frac{|DF_{pos}(T_i) - DF_{neg}(T_i)|}{\min(DF_{pos}(T_i), DF_{neg}(T_i)) \times TF(T_i, C_j)} \right) \quad (1)$$

Where, $DF_{pos}(T_i)$ and $DF_{neg}(T_i)$ are the number of positive and negative class of documents contain the term T_i respectively, $TF(T_i, C_j)$ is the occurrence count of term T_i in all the class C_j documents.

B. Existing Term Weight Measures

The researchers proposed various TWMs in different text classification based research domains. In this work, various TWMs are used for computing the term weight in vector representation of a document.

1) TFIDF (Term Frequency and Inverse Document Frequency)

TFIDF measure is a popular measure which is proposed to determine the term importance in a document [13]. TFIDF is

successfully used in text classification domain to calculate the term weight in a document. Equation (2) is used to compute the TFIDF of a term T_i in document D_k .

$$TFIDF(T_i, D_k) = TF(T_i, D_k) * \log\left(\frac{N}{1 + DF_i}\right) \quad (2)$$

Where, $TF(T_i, D_k)$ is the occurrence count of T_i in document D_k , N is documents count in dataset and DF_i is count of documents which contain term T_i at least one time.

2) TFIEF (Term Frequency and Inverse Exponential Frequency)

Equation (3) represents the TFIEF measure [14]. This measure uses the same information used in TFIDF measure.

$$TFIEF(T_i, D_k) = TF(T_i, D_k) * e^{-\frac{DF_i}{N}} \quad (3)$$

3) TFRF (Term Frequency and Relevance Frequency)

Equation (4) is used to determine the TFRF [15] of a term T_i in document D_k .

$$TFRF(T_i, D_k) = TF(T_i, D_k) * \log\left(2 + \frac{A}{C}\right) \quad (4)$$

Where, A and C are count of positive and negative class documents contain term T_i respectively.

4) TF-Prob

Equation (5) represents the TF-Prob measure [16].

$$TF - Prob(T_i, D_k) = TF(T_i, D_k) * \log\left(1 + \frac{A}{B} \cdot \frac{A}{C}\right) \quad (5)$$

Where, B is count of documents in positive class that doesn't contain term T_i .

5) TF-MI (Term Frequency – Mutual Information)

Mutual Information is one feature selection technique which calculates the term importance in a class of documents by considering the information of the way term was distributed in positive and negative class of documents. Deng et al., presented [17] a term weight measure based on mutual information. The term T_i weight in a document is computed by using TF-MI Equation (6).

$$TF - MI(T_i, D_k) = TF(T_i, D_k) * \log\left(\frac{AN}{(A+B) \times (A+C)}\right) \quad (6)$$

6) WLLR (Weighted Log Likelihood Ratio)

The WLLR term weight technique proved its efficiency in several text classification approaches [17]. The WLLR weight of term T_i in a document D_k that belongs to either positive or negative class C is determined by using Equation (7).

$$WLLR(T_i, D_k \in C) = \frac{A}{(A+B)} \log\left(\frac{A * (N - (A+B))}{C * (A+B)}\right) \quad (7)$$

7) TF-IDF-ICSDF (Term Frequency – Inverse Document Frequency – Inverse Class Space Density Frequency)

The IDF measure says that the terms which are discussed in less documents attained good weight. Like IDF measure, ICF (Inverse Class Frequency) measure says that the terms which are discussed less number of classes attained good weight. The TF-IDF-ICSDF measure was developed in the works of [18] by combining TF, IDF and ICSDF factors. The ICSDF is a variant of ICF, which gives the average number of classes that contain the given term. The ICSDF is determined by aggregating the probabilities of documents count in individual classes. The Equation (9) is used to calculate the term T_i weight using TF-IDF-ICSDF measure.

$$TF - IDF - ICSDF(T_i, D_k) = TF(T_i, D_k) * \left(\log\left(\frac{N}{DF(T_i)}\right) \right) * \left(\log\left(\frac{m}{\sum_{j=1}^m \left(\frac{n_{cj}(T_i)}{N_{cj}}\right)}\right) \right)$$

Where, m is classes count, $n_{cj}(T_i)$ is count of documents in class C_j that contain term T_i , N_{cj} is total documents count in class C_j .

8) Proposed Term Weight Measure (PTWM)

The TWM determines the term importance in a document. The Equation (10) used to calculate the proposed term weight measure.

$$PTWM(T_i, D_k \in C_j) = \frac{TF(T_i, D_k)}{TNTD_k} * \frac{TF(T_i, D_k \in C_j)}{TF(T_i, D_k \notin C_j)} * \frac{A+D}{B+C} * \left(\frac{A}{A+B} - \frac{C}{C+D} \right)$$

Where, A and C are number of documents that contain the term T_i in class C_j documents and other than class C_j documents respectively. B and D are number of documents that doesn't contain the term T_i in class C_j documents and other than class C_j documents respectively. $TF(T_i, D_k)$ is frequency of T_i in document D_k , $TNTD_k$ is total number of terms in document D_k . $TF(T_i, D_k \in C_j)$ is frequency of T_i in a class C_j of documents. $TF(T_i, D_k \notin C_j)$ is frequency of T_i in other than class C_j of documents.

The proposed term weight measure considers four different factors of information to determine the weight of the terms. The first factor says that the terms having more frequency in a document have more weight. The frequency of a term is normalized by dividing the term frequency with number of terms in a document.

The second factor is a ratio of number of times term T_i occurred in all documents of C_j class and all documents of other than C_j class. This factor gives more weight to the terms that are occurred more number of times in interested class and less number of times other than interested class.

The third factor becomes high when the $A+D$ value is higher and $B+C$ value is lower. The $A+D$ is higher when the term T_i occurred more documents of C_j class and less number of documents of other than C_j class. $B+C$ value is lower when the term T_i occurred more documents of C_j class and less number of documents of other than C_j class.

The fourth factor is the difference among number of documents of C_j class contains the term T_i and number of documents of other than C_j class contain the term T_i . $A+B$ is total number of documents in C_j class and $C+D$ is total documents count in other than C_j class. This factor says that when the term occurred in more C_j class documents than

other than Cj class documents then the difference is more and it implies the weight of Ti is more.

V. EXPERIMENTAL RESULTS

In this work, the experiment conducted with the top scored terms as features. The experiment started with top scored 2000 terms and incremented by 2000 terms in every next iteration. The experiment stopped at top scored 10000 terms. It was observed that the accuracy was dropped after experimenting with top scored 10000 terms. Two ML algorithms are used for producing the classification model. This model predicts the accuracies of age and gender prediction. The PAN 2014 competition reviews dataset was considered for experimentation. The gender dataset contains two classes such as male and female, age dataset contains five classes such as 18-24, 25-34, 35-49, 50-64 and 65+. Table 2 shows the SVM classifier accuracies of gender prediction when experiment conducted with different number of terms and various term weight measures.

TABLE 2. GENDER PREDICTION RESULTS OF SVM CLASSIFIER ON REVIEWS DATASET

Features / TWM's	TF-IDF	TFI EF	TFR F	TF-Pro b	TF-MI	WL LR	TF-IDF-ICS DF	PST WM
2000	0.62 94	0.63 23	0.70 25	0.71 36	0.74 11	0.74 44	0.73 61	0.805 3
4000	0.64 61	0.65 20	0.72 85	0.73 15	0.75 32	0.76 35	0.70 15	0.821 5
6000	0.65 15	0.67 23	0.73 35	0.74 12	0.76 74	0.77 97	0.75 64	0.839 7
8000	0.66 65	0.68 48	0.75 35	0.76 13	0.77 35	0.79 44	0.64 14	0.841 9
10000	0.68 48	0.69 51	0.76 20	0.77 87	0.79 74	0.80 79	0.83 61	0.853 7

In Table 2, the proposed TWM attained best accuracy of 85.37% for gender prediction when compared with other TWMs. It was recognized that the gender prediction accuracy was enhanced when the terms count is increased for document vector representation. Table 3 shows the SVM classifier accuracies of age prediction when experiment conducted with different number of terms and various term weight measures.

TABLE 3. AGE PREDICTION RESULTS OF SVM CLASSIFIER ON REVIEWS DATASET

Features / TWM's	TF-IDF	TFI EF	TFR F	TF-Pro b	TF-MI	WL LR	TF-IDF-ICS DF	PST WM
2000	0.60 61	0.61 92	0.66 61	0.67 97	0.69 61	0.72 74	0.73 05	0.740 1
4000	0.61 17	0.62 15	0.67 17	0.69 90	0.70 31	0.73 61	0.74 25	0.758 9
6000	0.62 92	0.64 23	0.69 92	0.70 72	0.72 82	0.74 61	0.75 15	0.761 2
8000	0.64 53	0.65 15	0.70 53	0.71 01	0.73 34	0.75 34	0.76 91	0.782 3
10000	0.65 15	0.67 25	0.72 15	0.73 21	0.75 21	0.76 67	0.78 65	0.796 3

In Table 3, the proposed TWM attained best accuracy of 79.63% for prediction of age when compared with other TWMs. It was recognized that the age prediction accuracy was enhanced when the terms count is increased for document vector representation. Table 4 shows the RF classifier accuracies of gender prediction when experiment conducted with different number of terms and various term weight measures.

TABLE 4. GENDER PREDICTION RESULTS OF RF CLASSIFIER ON REVIEWS DATASET

Features / TWM's	TF-IDF	TFI EF	TFR F	TF-Pro b	TF-MI	WL LR	TF-IDF-ICS DF	PST WM
2000	0.65 02	0.66 63	0.73 92	0.74 65	0.75 92	0.79 61	0.80 97	0.832 3
4000	0.67 34	0.67 79	0.74 57	0.76 19	0.77 05	0.80 11	0.81 72	0.842 2
6000	0.67 92	0.68 09	0.76 07	0.77 61	0.78 62	0.81 16	0.83 65	0.847 6
8000	0.68 45	0.70 15	0.77 45	0.79 03	0.80 93	0.81 97	0.84 93	0.863 5
10000	0.70 61	0.71 97	0.78 14	0.80 01	0.81 41	0.83 23	0.85 28	0.879 6

In Table 4, the proposed TWM attained best accuracy of 87.96% for prediction of gender when compared with other TWMs. It was recognized that the gender prediction accuracy was enhanced when the terms count is increased for document vector representation. Table 5 shows the RF classifier accuracies of age prediction when experiment conducted with different number of terms and various term weight measures.

In Table 5, the proposed TWM attained best accuracy of 82.58% for age prediction when compared with other TWMs. It was recognized that the age prediction accuracy was enhanced when the terms count is increased for document vector representation

TABLE 5. AGE PREDICTION RESULTS OF RF CLASSIFIER ON REVIEWS DATASET

Features / TWM's	TF-IDF	TFI EF	TFR F	TF-Pro b	TF-MI	WL LR	TF-IDF-ICS DF	PST WM
2000	0.61 61	0.64 55	0.69 65	0.71 02	0.71 71	0.74 00	0.76 81	0.781 6
4000	0.62 20	0.65 24	0.70 66	0.72 23	0.73 80	0.75 12	0.77 93	0.791 7
6000	0.64 66	0.67 93	0.71 45	0.74 48	0.74 90	0.76 89	0.78 09	0.809 2
8000	0.65 92	0.68 23	0.73 38	0.75 10	0.75 29	0.78 23	0.80 74	0.815 3
10000	0.67 76	0.69 61	0.74 67	0.76 53	0.77 90	0.79 51	0.81 08	0.825 8

VI. DISCUSSION ON RESULTS

In this work the experiment carried out with two machine learning algorithms such as SVM and RF for gender and age prediction. Different term weight measures are used to determine the importance of term in a document. The Fig. 2 shows the accuracies of SVM and RF classifiers for gender prediction.

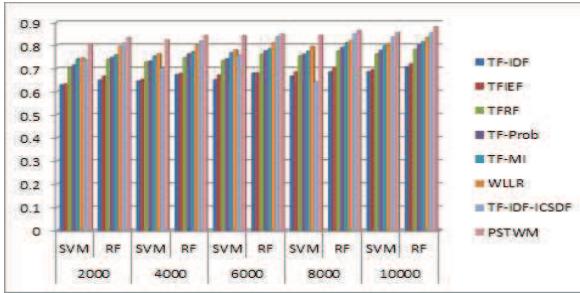


Fig. 2. The Accuracies Gender prediction

In Fig. 2, it was observed that the proposed term weight measure attained best accuracy for gender prediction when experimented with top scored 10000 terms. The RF classifier obtained good accuracies for gender prediction than the accuracies of SVM classifier. The Fig. 3 shows the accuracies of SVM and RF classifiers for gender prediction.

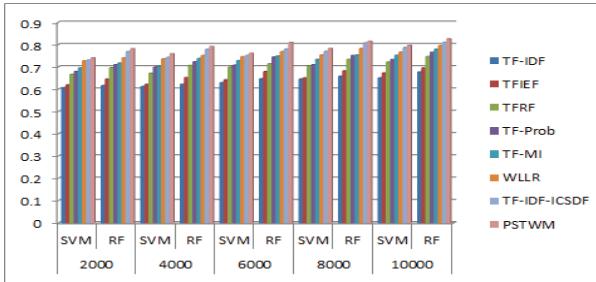


Fig. 3. The Accuracies of Age Prediction

In Fig. 3, it was observed that the proposed term weight measure attained best accuracy for age prediction when experimented with top scored 10000 terms as features. The RF classifier obtained good accuracies for age prediction than the accuracies of SVM classifier.

VII. CONCLUSIONS AND FUTURE SCOPE

Author profiling is defined as the task of identifying one or more attributes such as gender, age, personality traits of an author based on how they write. In this work, the age and gender was predicted from the dataset of PAN competition 2014 reviews. The term weight measure based approach was proposed to predict the author attributes of age and gender. In this approach, a new TWM was proposed to determine the weight of a term in the document vector representation. The content based features of terms that are selected by the feature selection algorithm are used as features to represent the document vectors. In vector representation, the term value is computed by using term weight measures. The experiment conducted with different TWMs and proposed TWM. Two ML algorithms such as SVM and RF are used to develop the classification model. The RF classifier performance is good than SVM classification algorithm. The proposed term weight measures attained best accuracies of 0.8796 and 0.8258 for gender and age prediction respectively than the accuracies of other TWMs when RF classifier is used. The results are good when compared with

other approaches in AP for predicting gender and age of author.

In future work, we are planning to implement a new document representation technique with a combination of new feature selection technique and new term weight measure. We are also planning to implement deep learning techniques to predict the accuracies of gender and age prediction.

REFERENCES

- [1] Koppel M, Argamon S, Shimoni A. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*; 2003. p. 401–12.
- [2] E. Stamatatos. 2009. A Survey of Modern Authorship Attribution Methods. *JASIST*.
- [3] Raghunatha Reddy, T., Vishnu Vardhan, B., Vijayalal Reddy, P., A survey on author profiling techniques. *Int. J. Appl. Eng. Res.* 11(5), 3092–3102 (2016).
- [4] 4 Muhammad Waqas Anjum Ch, Waqas Arshad Cheema, “A Study of Content Based Methods for Author Profiling in Multiple Genres”, International Journal of Scientific & Engineering Research Volume 9, Issue 9, September-2018, PP. 322-327
- [5] 5 Pushkar Mishra, Marco Del Tredici, Helen Yannakoudakis, Ekaterina Shutova, “Author Profiling for Hate Speech Detection”, arXiv:1902.06734v1 [cs.CL] 14 Feb 2019
- [6] 6 Chiyu Zhang and Muhammad Abdul-Mageed, “BERT-Based Arabic Social Media Author Profiling”, arXiv:1909.04181v3 [cs.CL] 31 Oct 2019
- [7] 7 Zhang, C., & Abdul-Mageed, M. (2019). BERT-Based Arabic Social Media Author Profiling. In: Mehta P., Rosso P., Majumder P., Mitra M. (Eds.) Working Notes of the Forum for Information Retrieval Evaluation (FIRE 2019). CEUR Workshop Proceedings. CEUR-WS.org, Kolkata, India.
- [8] 8 Inna Vogel and Peter Jiang, “Bot and Gender Identification in Twitter using Word and Character N-Grams”, CLEF 2019, 9-12 September 2019, Lugano, Switzerland.
- [9] 9 Kowsari, K., Heidarysafa, M., Odukoya, T., Potter, P., Barnes, L. E., & Brown, D. E. (2020). Gender detection on social networks using ensemble deep learning. arXiv preprint arXiv:2004.06518.
- [10] 10 Moniek Nieuwenhuis and Jeroen Wilkens, “Twitter Text and Image Gender Classification with a Logistic Regression N-gram Model”, Notebook for PAN at CLEF 2018.
- [11] 11 <https://pan.webis.de/clef14/pan14-web/author-profiling.html>
- [12] 12 Rehman, A. , Javed, K. , Babri, H. A. , & Saeed, M. (2015). Relative discrimination criterion-A novel feature ranking method for text data. *Expert Systems with Applications*, 42 , 3670–3681.
- [13] 13 Debole, F., & Sebastiani, F. (2003). Supervised term weighting for automated text categorization. In *Proceedings of the 2003 ACM symposium on applied computing* (pp. 784–788). Melbourne, Florida, USA.
- [14] 14 Z. Tang, W.Q. Li, Y. Li, An improved term weighting scheme for text classification, *Concurr. Comput.: Pract. Exper.* 32 (2020) e5604.
- [15] 15 M. Lan, C. Tan, J. Su, Y. Lu, Supervised and traditional term weighting methods for automatic text categorization, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (4) (2009) 721–735.
- [16] 16 Liu, Y., Loh, H. T., & Sun, A. (2009). Imbalanced text classification: A term weighting approach. *Expert Systems with Applications*, 36 (1), 690–701. <http://doi.org/10.1016/j.eswa.2007.10.042>
- [17] 17 Z. Deng, K. Luo, H.Yu, A study of supervised term weighting scheme for sentiment analysis, *Expert Syst. Appl.* 41(2014) 3506–3513.
- [18] 18 Ren, F., & Sohrab, M. G. (2013). Class-indexing-based term weighting for automatic text classification. *Information Sciences*, 236 , 109–125. <http://doi.org/10.1016/j.ins.2013.02.029>.

Joint Spectral-Spatial Feature Using Deep 3-D CNN for Hyperspectral Images

1st M. Kavitha

Department of Electronics and
Communication Engineering,
Sri Venkateshwara College of
Engineering,
Chennai India,
kavithamunisamy@gmail.com

2nd R. Gayathri

Department of Electronics
and Communication Engineering,
Sri Venkateshwara College of
Engineering,
Chennai India,
rgayathri@svce.ac.in

Abstract— The importance of hyperspectral imaging has been extensively adopted to variety of real time applications. HSIC classification (HSI) is a difficult process due to the large amount of similarity between classes, variability within classes, overlap, and nested regions. Deep learning techniques are highly recommended for improvised HSI classification, with convolution neural network (CNN) being at most prominent approach to classifying both Spectral-Spatial information. Because of the large spectral dimensions and bands, a 3D-CNN is recommended for complex data computation. The extracted features' similarity features must perform better. As a result, the paper discusses, an enhanced 3D-CNN model is proposed that joints both spectral and spatial features to improve performance. First, the CNN Layers are increased and normalized, and then Spectral-Spatial information is extracted thoroughly using Principal Component Analysis (PCA). Second, spectral and spatial data are combined to create a joint feature map. Finally, 3-D CNN is used in the classification process. To classify and validate the vegetation information, standard HSI datasets (Indian Pines and Salinas) are used. The attained results of the existing method are compared to the proposed work.

Keywords— Convolution Neural Network (CNN), Hyperspectral Image (HSI), Classification, spectral-spatial features.

I. INTRODUCTION

Various approaches to hyperspectral image (HSI) classification have been developed over the last few decades. Deep learning approaches based on convolutional neural networks (CNNs) have recently made significant progress in computer vision, showing strong efficiency in image processing. As a result, research into CNN models has gained a lot of interest, which has led to the application of CNN in numerous subsets of image processing, such as remote sensing image processing. One of the highlights in the remote sensing world has long been hyperspectral image classification. Deep learning-based methods, have grown in popularity and have shown impressive results. Various CNN-based deep learning models have been applied to HSI classification with limited labelled samples, because CNN can extract low, mid, and high-level spatial characteristics. Different ways have been proposed to either expand the training set or minimise the parameters of the network in order to appropriately train CNN in the context of limited labelled samples and fine-tuned hyperparameters. The ImageNet dataset, which is a

large collection of realistic images, yielded good classification results using a deep learning model called AlexNet [18]. Since then, a never-ending stream of novel networks has evolved, constantly motivating the feature extraction and reuse concept. For the first time in 2014 deep learning approach for remote sensing image study was used in paper [1] stating its applications. This article used Principal Component Analysis (PCA), Logistic Regression, and Stacked Autoencoder as deep learning approaches. A 5-layer CNN [2] with an impressive result were classified efficiently. Using CNN morphological profiles [3] were extracted prominently. A 3D CNN with various filter sizes [4] were casted to feature extract from both spectral and spatial data, which obtained a significant results of classification. The joint features of spectral and spatial information [5], using scattering wavelet and Gabor filter were used to extract features to perform classification. Traditional feature extraction methods, on the other hand, are based on learning algorithm and handcrafted attributes that rely heavily on contextual information [6]. As a result, Deep learning methods and models have been suggested to address the existing challenges, i.e., to automatically study the required features from raw HSI data, and have achieved an enormous achievement for Hyperspectral Image Classification (HSIC).

2D CNN models [7] showed an advancement improvement in extracting spectral-spatial features for HSIC. Although, 2D CNN was developed to extract features individually but avoided the joint spectral-spatial features. Thus, this motivated the research to explore the joint features in HSIC. A 3-D CNN with increased number of layers is proposed for HSIC. Initially, data normalization is performed without losing any information from the dataset. Then, integration of spectral-spatial feature is performed using PCA and 3-D CNN. Authors [16] built a 3D-CNN model with 3D convolutional layers and pooling layers to improve classification efficiency by delving into spatial-spectral information in depth. Deeper networks allow for more robust and detailed characteristics, as well as a more complex network design. In paper [17], authors conducted extensive trials with various numbers of training samples and discovered that the CNN model degrades frequently as the sample size decreases. Finally, to obtain better classification dropout and batch normalization are used. The main contribution of the proposed work is listed below:

- Data normalisation is performed on the dataset without losing information, as it reduces space

requirements. This data standardisation aids in the reduction of computing load.

- The joint integration of features from spectral and spatial data creates a new HSI classification paradigm. This model is constructed using a combination of Principal Component Analysis and Convolutional Neural Network techniques (CNN) respectively.
- Even with minimal training data, the 3D-CNN proposed HSI classification using deep CNN exhibits decent overall classification accuracy. To improve accuracy, dropout and batch normalisation (BN) were also used.

The remainder of the proposed paper is organised as follows: in section II discusses the implemented methodology. The experimental benchmark Datasets, their Results, and Discussion about them are described in Section III. Finally, Section IV of the article closes with potential future use for the research directions.

II. PROPOSED METHOD

The methodology used in the proposed work for remote sensing HSI classification as shown in Figure 1. The HSI dataset initially examined and standardised. filtering on the principal components is used to extract spatial characteristics. These spatial and spectral features are joint together. The huge hyperspectral data with conjoined features is basically split into training and testing groups and given into a CNN architecture.

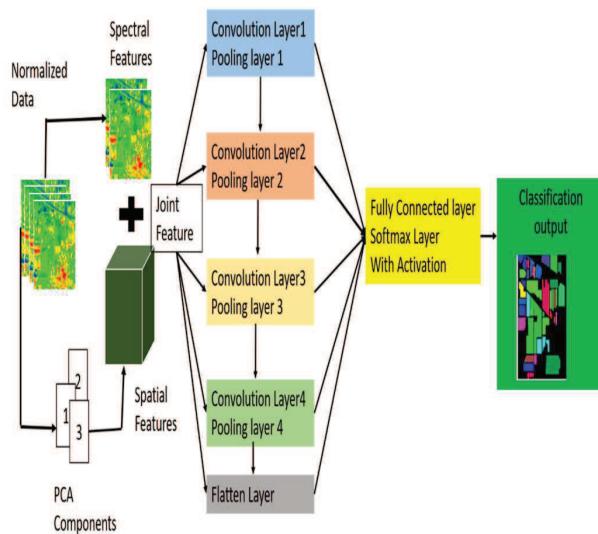


Figure 1 Proposed enhanced 3D-CNN Model.

A good classification model must take into account that HSI pixels have significant interclass similarity, intra-class variability, and stacked regions [8-9]. Hence, to overcome these issues Principal Component Analysis (PCA) process is

to reduce images for each band without disturbing the spatial dimensions of the dataset.

CNN model comprises of convolution (C_1, C_2 and C_3) and max-pooling (P_1, P_2 and P_3) operations with different filter size of ($f_1 = 32, f_2 = 64$ and $f_3 = 128$). First convolution layer takes the input as ($M_1, N_1, 1$) after the convolution operation the with (f_1) filter the output is (M_1, N_1 and f_1) and (M_2, N_2 and f_1) and in second layer the output feature has (M_2, N_2 and f_2) and (M_3, N_3 and f_2). Lastly, the output feature has (M_3, N_3 and f_3). to this extracted output flatten operation is performed flatten layer to ensure that the classifier can differentiate spatial information throughout multiple spectral bands without losing any information. At fully connected layer the feature size is reduced. With help of soft-max operation the feature is classified.

Table-1 The model summary of the proposed 3D-CNN model with knowing window size of (9 X 9)

Operating Layers	Output Feature	No. of Parameter
Input Layer	(9, 9, 20, 1)	0
Conv3D -1	(7, 7, 65, 64)	2944
Conv3D -2	(5, 5, 17, 32)	92192
Conv3D -3	(3, 3, 2, 32)	46112
Flatten-1	576	0
Dense-1	128	73856
Dropout-1	576	0
Dense-2	128	16512
Dropout-2	128	0
Dense-3	16 classes	2064
Total obtained trainable Parameters = 2,33,680		

Table 1 contains further information on the proposed model. 3D CNN model has a total of 233,680 trainable parameters. The weights are generated at first, then optimised via Adam optimizer and a soft-max regression loss function. Without batch normalisation the usual weights are rationalized with 256 mini-batch sizes for 50 epochs. Overfitting is a critical challenge in HSI data as it has inadequate availability of training-set samples. Regularization solves the problem of overfitting. Dropout method [10] is used in 3D-CNN model it helps in removing unwanted nodes, thus model has a unique node set. Batch normalisation is used [11] to increase the performance of the proposed 3D-CNN model. The data flow should be very ease and data should be distributed normally else it is hard to train the model. As a result, batch normalisation is utilized after each layer to make training go smoothly. Figure 2 shows the

sample training and kernel validation loss and accuracy of the proposed 3D-CNN system on Indian Pines dataset.

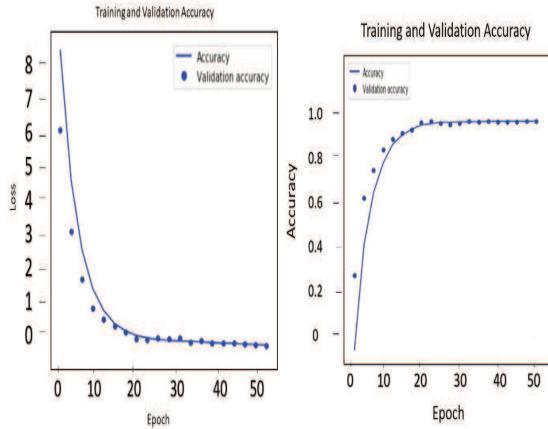


Figure. 2: Loss and Accuracy for samples Training and kernel Validation sets on Indian Pines Dataset with 9x9 window

III. EXPERIMENT RESULTS AND DISCUSSIONS

Datasets

The Salinas dataset (SA) was gathered using an AVIRIS sensor of Salinas Valley, California. SA is 512 x 217x 224 pixels in size, 3.7 meter spatial-resolution and 512 x 217 spatial and 224 spectral data dimensions. SA is made up of vineyards, veggies, and bare dirt. SA is shown in Figure 3 and its divided into 16 classes. Before analysis, the noisy water absorption bands 108 112, 154 167, and 224 are excluded. The Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) sensor produced the Indian Pines Dataset (IP), northwest Indiana. IP has a size of 145x145x224 in the wavelength range of 0.4 2.5 x10-6 metres, with 145x 145 being the spatial dimension and 224 being the spectral dimension. IP is made up of 1-3rd woodland, 2-3rd farmland, and rest are evergreen vegetation. The IP illustration is shown Figure 4 along with ground truth which consists of 16 classes. the noisy water absorption band have been dissented. Thus, 200 bands are utilized.

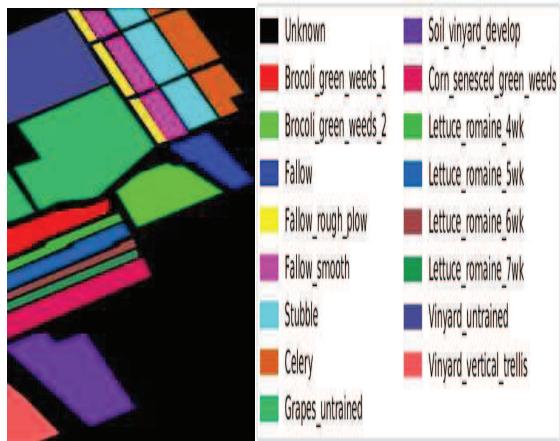


Figure 3 Salinas dataset with color codes

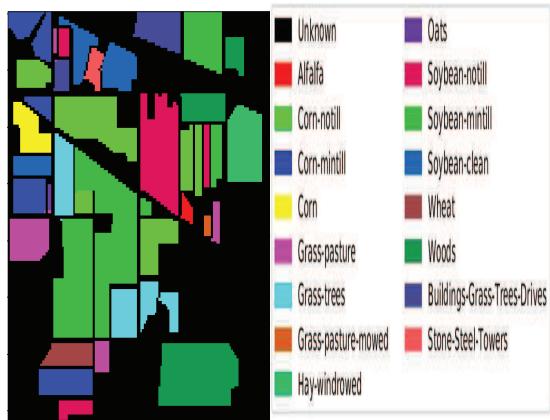


Figure (4) Indian Pines dataset with color codes

The experiments were carried out on Google Colab [12], an online platform. Google Colab is an available online platform that runs in any environment and requires a fast internet connection. For data calculation, Google Colab offers the option of running the programs on a Python 3 jupyter notebook which has Graphical Processing Unit (GPU), 32 GB of RAM. The Test and Train set is split into a 70-30 percent ratio. The learning rate is fixed to 0.001, ReLU is used as an activation function in all layer other than Softmax- layer.

The confusion matrices were used to calculate the dataset Average Accuracy (AA), class-wise Overall Accuracy (OA), and Kappa (κ) coefficients for evaluation purposes. The average class category classification performance can be determined by AA, and the number of accurately predicted class samples out of the entire test samples is represented by OA. Finally, is an arithmetic metric that takes into account shared information about understanding between classification and provided ground-truth maps.

The training and validation accuracy and loss of proposed 3DCNN model for a 50 epochs is elaborated in Figure 4. The above figures constitute the proposed model is congregated near to 41st epoch. Moreover, the classification feature maps are shown in Figure 5 -6 with a window size of 9X9. The

proposed models performance is compared with existing classifiers. In the Table 2 the performance accuracy of different classifiers is shown (SVM, 1-D CNN, 2-D CNN, Proposed-3D CNN) [13-14]

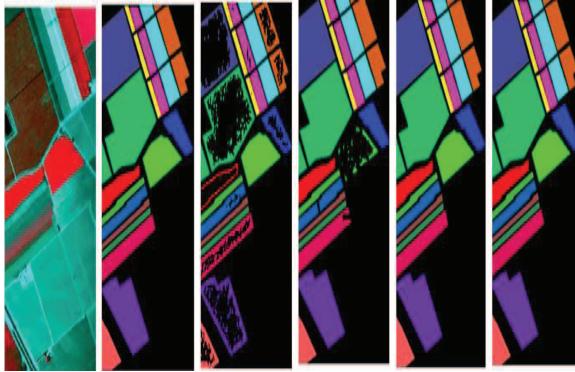


Figure 5 Salinas Classification maps (a) RGB Image (b) Ground Truth (c) SVM (d)1-D CNN (e) 2-DCNN (f) Proposed 3-D CNN

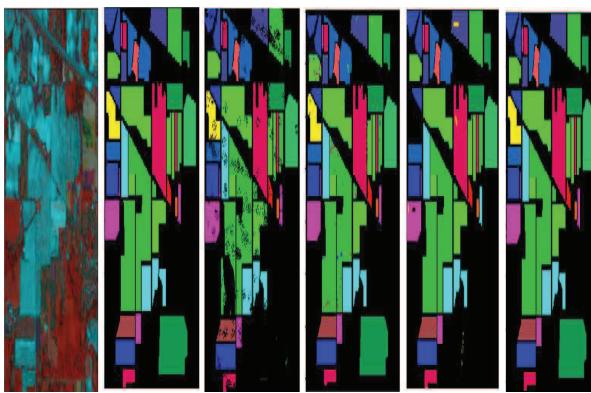


Figure 6 Indian Pines Classification maps (a) RGB Image (b) Ground Truth (c) SVM (d)1-D CNN (e) 2-DCNN (f) Proposed 3-D CNN

Table 2 Comparative performance evaluations with 15% of training samples.

Da tas et	SVM		1-D CNN		2-D CNN		Proposed 3-D CNN	
	OA	AA	OA	AA	OA	AA	OA	AA
SA	87 .1 4	88 .3 6	84 .7 0	87 .5 6	87 .1 7	87 .3 7	99 .8 0	99 .6 4
IP	83 .4 6	83 .5 1	81 .9 7	82 .6 9	81 .8 0	82 .3 5	84 .7 9	81 .2 6

IV. CONCULSION

The most challenging task in remote sensing is hyperspectral Image (HSI) classification because of the significant class similarity and variability. The proposed enhanced 3-D CNN designed model has a PCA normalised feature input to extracted features of both spatial and spectral information. The dataset image to feature extracted image is joint and fed to proposed 3-D CNN model. Batch normalization and Dropout regularization strategies are also

used. The obtained classification results on benchmark datasets are much improvised results than the traditional approaches. In future, the usage of proposed 3-D CNN model on different HSI datasets need to be carried out.

REFERENCES

- [1] 1. Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep learning-based classification of hyperspectral data, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, Jun. 7 (2014) 2094–2107.
- [2] 2. W. Hu, Y. Huang, L. Wei, F. Zhang, H. Li, Deep convolutional neural networks for hyperspectral image classification, *J. Sensors* 2015(Jan.) (2015) 258619
- [3] 3. E. Aptoula, M.C. Ozdemir, B. Yanikoglu, Deep learning with attribute profiles for hyperspectral image classification, *IEEE Geosci. Remote Sens. Lett.* 13 (12) (2016) 1970–1974
- [4] 4. H. Lee, H. Kwon, Contextual deep CNN based hyperspectral classification, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, 2016, pp. 3322–3325. <https://doi.org/10.1109/IGARSS.2016.7729859>.
- [5] 5. Y. Zhong, A. Ma, and L. Zhang, “An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1235–1248, 2014.
- [6] 6. S. Roy, G. Krishna, S. R. Dubey, and B. Chaudhuri, “Hybridsn: Exploring 3-d-2-d cnn feature hierarchy for hyperspectral image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 277–281, 06 2019.
- [7] 7. L. Huang and Y. Chen, “Dual-path siamese cnn for hyperspectral image classification with limited training samples,” *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020
- [8] 8. M. Ahmad, D. Ihsan, and D. Ulhaq, “Linear unmixing and target detection of hyperspectral imagery using osp,” 01 2011, pp. 179–183.
- [9] 9. M. Ahmad, S. Lee, D. Ulhaq, and Q. Mushtaq, “Hyperspectral remote sensing: Dimensional reduction and end member extraction,” *International Journal of Soft Computing and Engineering (IJSC)*, vol. 2, pp. 2231–2307, 05 2012.
- [10] 10. G. Hinton, et al., Improving neural networks by preventing co-adaptation of feature detectors, *Comput. Sci.* 3 (4) (2012) 212–223.
- [11] 11. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, Unpublished paper, 2015. [Online]. Available <https://arxiv.org/abs/1502.03167>.
- [12] 12. T. Carneiro, R. V. M. Da Nobrega, T. Nepomuceno, G.-B. Bian, V. H. C. De Albuquerque, and P. P. Reboucas Filho, “Performance analysis of google colabatory as a tool for accelerating deep learning applications,” *IEEE Access*, vol. 6, pp. 61 677–61 685, 2018.
- [13] 13. L.N.P. Boggavarapu, Manoharan Prabukumar, Classification of Hyper Spectral Remote Sensing Imagery Using Intrinsic Parameter Estimation, Springer Nature Switzerland AG 2020, ISDA 2018, AISC 941, pp. 1–11, 2020. https://doi.org/10.1007/978-3-030-16660-1_83
- [14] 14. M. He, B. Li, and H. Chen, “Multi-scale 3d deep convolutional neural network for hyperspectral image classification,” in 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 3904–3908
- [15] 15. Shrutiika S. Sawant, M. Prabukumar, Unsupervised band selection based on weighted information entropy and 3D discrete cosine transform for hyperspectral image classification, *Int. J. Remote Sens.* 41 (10) (2020) 3948–3969. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].
- [16] Chen, Y.S.; Jiang, H.L.; Li, C.Y.; Jia, X.P.; Ghamisi, P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 2016, 54, 6232–6251.
- [17] Feng, F.; Wang, S.; Wang, C.; Zhang, J. Learning Deep Hierarchical Spatial–Spectral Features for Hyperspectral Image Classification Based on Residual 3D-2D CNN. *Sensors* 2019, 19, 5276
- [18] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th

Traffic Violation Invigilation using Transfer Learning

N.R. Rajalakshmi

Computer Science & Engineering

Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and
Technology, Tamilnadu, India

K.Saravanan,

Computer Science and Engineering

Anna University Regional Campus

Tirunelveli Tamilnadu, India

Saravanan.krishnan@auttvl.ac.in

Abstract— The modern era of evolution and urbanization in this populated world makes the abundance of vehicles on the roads. The plenty number of road accidents are increased as a consequence of violation of traffic rules. Even though many techniques are applied to invigilate the traffic violation, still it is difficult to supervise the traffic violation. An automated system for traffic invigilation is required to handle and track the people who are infringing the rules. By using Convolutional Neural Network (CNN), the right model is built to predict the class of violation. In this work, the classification of the vehicle, detection of helmet violation, wearing mask is also done using CNN model. Transfer Learning is applied here to use the learned weightage from the existing model to improve the detection.

Keywords— Regional Convolutional Neural Network, AlexNet, VGG16, SSD, Mobile Net, Object Detection, Feature Map, Convolutional Layer

I. Introduction

In today's world, the number of cars and other vehicles has increased dramatically over the previous few periods. So, the density of traffic is increasing, and traffic offences are increasing day by day. Increased traffic violations will become a major cause of accidents and property destruction, potentially resulting in serious injuries and death. 9 % accidents account in Germany and other European countries, 20% in Australia, New Zealand, and other Pacific countries, and a 34% in emerging countries like India, Pakistan, and other Asian countries, according to the census. The traffic offences have a negative impact on society, which are a huge problem. [1] A significant number of automobiles causes traffic congestion in places like toll booths, malls and parking lots. Road accidents, particularly those involving motorcycles, are one of the leading causes of unusual deaths [3,4, 5]. The traffic department is working, but keeping track of each and every vehicle for traffic management and law enforcement purposes has become very difficult. As a result, the number of violations has increased. The department can handle some issues, but they are unable to focus on many violations. The goal is to automate the detection of traffic signal violations and make it easier for the traffic police department to monitor the traffic. The input of automation

system includes video streams and it may check for any violations committed by the vehicle's driver as well as identify the drivers' personal information. Helmet is safety gear that gives protection to motorcyclists. Because, it is an essential one, It gives the potential to improve compliance. The public is affected either directly or indirectly as a result of crimes such as vehicle's speed and not wearing helmets and masks. This could be remedied by doing monitoring the motorcyclists and analysing the image to quickly identify the infractions. As a result, human lives are saved. The front-end elements of the car, such as the contour, windshield, rear-view mirror, and licence plate, can be used to classify the vehicle. The detection of traffic violations is accurate one if it has been done under the deep learning and computer vision umbrella. Deep learning technology used to accurately detect the vehicle with high accuracy and its activity to accomplish this automation as shown in fig.1. It detects whether a vehicle has been involved in a violation and then checks for the vehicle's details and responds to the user accordingly. The object detection and localization in images and videos is accurately done by emerging popular convolutional neural network method. By using CNN, the road vehicles like motorcycles, buses, minibuses, trucks, cars, and vans has been classified for vehicle analysis in traffic violation.

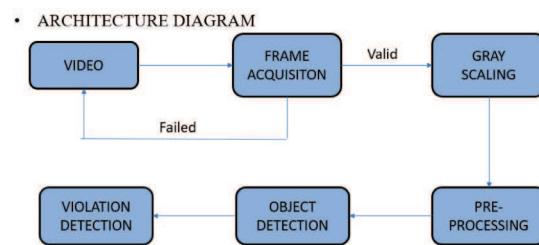


Fig.1. Block Diagram of Object Detection

II. Literature Survey

[5] They are just that have been presented a few of the R-CNN-based models, such as Region with CNN features (R-CNN) detection framework, Quick R-CNN, FCNN, SPP-net. Object detection may be accomplished using a variety of deep learning algorithms, including R-CNN and YOLO. The

large convolutional neural network R-CNN use the Object detector function to finds the relevant regions from 2000 candidate region using a search selective strategy as shown in fig.4. Finally, the extracted features from the candidate box belong to a specific class is classified by using Support Vector Machine. After classification of objects, the regions are filtered out by using greedy non-maximum suppression algorithm. The discrepancies between the anticipated and true bounding boxes are calculated. RCNN takes a long time to extract the candidates frames and the training data storage space also considerable one [3]. As a result, Faster RCNN shares the convolutional feature with detection network to brings down the region proposal time. YOLO Convolutional Neural Network are also used for object detection. The main task of this is, object categorization, or identifying the specific objects in a frame. Object detection is a more difficult problem, which is a combination of classification and localization. The customized object detector recognizes the items in the input. YOLO CNN predicts the object by making bounding boxes and estimating associated confidence scores. Non max suppression is applied to avoid concurrencies or redundant bounding box co-ordinates. The region of interest is marked, and after which the class probabilities and bounding box co-ordinates of the specific video are calculated. If any car crosses that area of interest, it is considered as a violation, and a snapshot of that vehicle is taken and saved locally. As a result, detection of violations is made easier. YOLOv3 is a one of the networks, has a characteristic of detection of small target. It uses unique logistic classifiers for each class, which is used to create multi-label classifications. Every bounding box forecasts the classes, and the bounding box may be used for multi-label classifications. YOLOv3 can make reliable predictions at various scales and uses logistic regression to estimate the objectiveness score. Based on score value, it will forecast only one bounding box in advance for one ground truth item. If objectiveness score that is higher than the threshold but lower than the best, both classification and detection loss will be occurred. [9] introduced a SSD (Single Shot MultiBox Detector) technique which has higher detection accuracy and faster speed by combining the Faster R-CNN with the YOLO. [19][20][21].

III. SYSTEM ARCHITECTURE

The vehicles can be easily identified by their licence plate numbers using OCR (Optical Character Recognition). Additionally, licence plate numbers are useful in ascertaining the details of the motorist, which can aid in identifying traffic rule violators. The driver can also be blamed under the traffic rule violations based on driving speed. The distance between frames of a video and the video length aids to compute the vehicle's speed. The violation can be detected if the speed is more than threshold value. But, still the manual operations are carried out in this areas. To make these procedures easier and more efficient, a system is needed to recognise the violation of vehicle. On road controls, traffic control devices are mounted to guarantee that vehicles do not infringe traffic

rules and regulations. Traffic control equipment, such as cameras, are installed on road to ensure and check the violations of traffic rules. By using this camera's input frame, the classification of the vehicle, detection of helmet violation, wearing mask is done using pretrained model like SSD Mobile net, Alex Net, VGG16. This model has a lot of positive effects and aids in society's proper functioning. Alex Net is a CNN model that gives high accuracies for vehicle categorization. It is a pioneering object-detection architecture and gives better performance with convolutional layers. It comprises of totally 23 layers which are convolutional, max pooling, fully connected layers and probabilistic SoftMax layer for classification of 1000 categories. A kernel or filter that slides across the input image to extract the feature map in the object during the convolution process. The activation function ReLU is used in Alex Net. The 25 % error rate roughly occurs in ReLU activation network, But, ReLU helps with the vanishing gradient problem. Alex Net implemented Local Response Normalization to do a normalisation in a neighbourhood of pixels. It takes an ImageNet image to train the network, which is primarily an RGB image with a size of 227x227 pixels. It solves the classification problem in which the input image may be an image of a class of car, motorcycle, truck, bus and others, and the output is a vector of one thousand integers. As a result, the sum of each member in the output array is 1.

The batch RGB images with a size of 227x227x3 is taken as input and returns probability vector of a 1000x1, one for each class. Max pool layers 3*3 with strides of 2 smaller than the window size are referred as overlapping max pool layers. However, the overlapped max-pooling layer is used in Alex Net after the convolution layers detects the drop in error, the overlapped pooling makes overfit. By making mirroring and resizing the images, overfitting is avoided and improve the variance in the dataset.

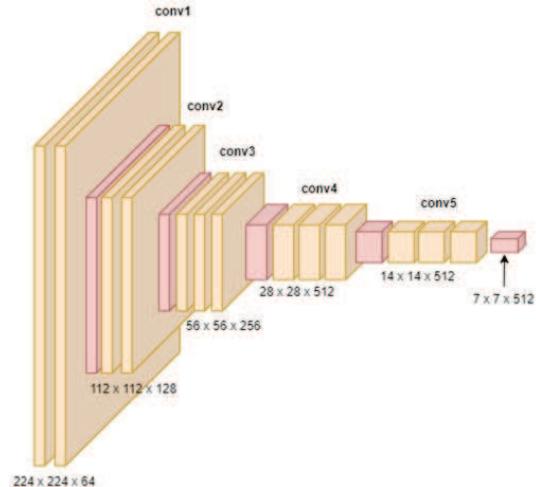


Fig.2. Base Model using Alex Net

Transfer learning is applied to utilize the base model of pretrained convolutional neural network (Alex Net) to recognise an object as shown in fig.2. Transfer learning makes considerably faster and easier for the object identification & classification. It fine tune the trained network rather than training a network with randomly chosen weights. Later, the network learns and develops rich feature representations for a variety of images.



Fig.3.Examples of Car, Motorbikes and Buses on Traffic

Convolutional layers with pooling are added additionally to replace or optimize the few top layers of the Alex Net model, which is shown in fig.5. The top layers are altered in Alex Net model. As a result, the top layers of Alex Net are suggested to be replaced with Convolutional layers, Max pooling layer. The Alex network is used here to utilise the learnt characteristics of motorbikes, vehicles, and buses collected from the enlarged dataset. The convolutional filter in each extra layer makes a fixed number of predictions. Finally, one Soft max classifier is used for 6 label classification. The Soft Max layer is added to handle the six classes of motorbikes, motorcycles, pedestrian, car, truck and buses. The classification of the vehicle which is shown in fig.4 & 6 is done by using the proposed RCNN model

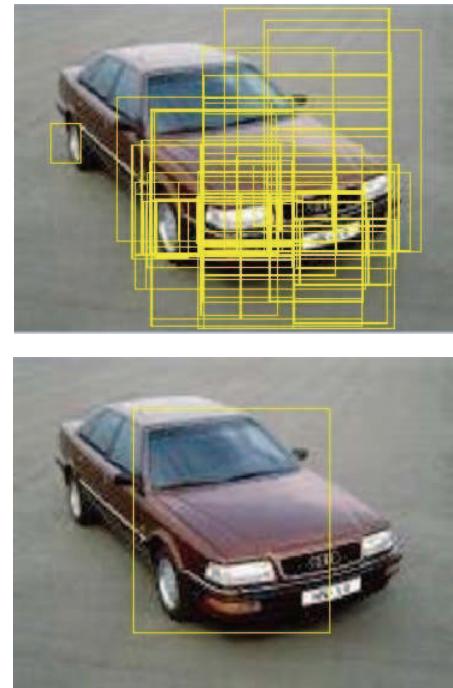


Fig. 4. Object detection using RCNN

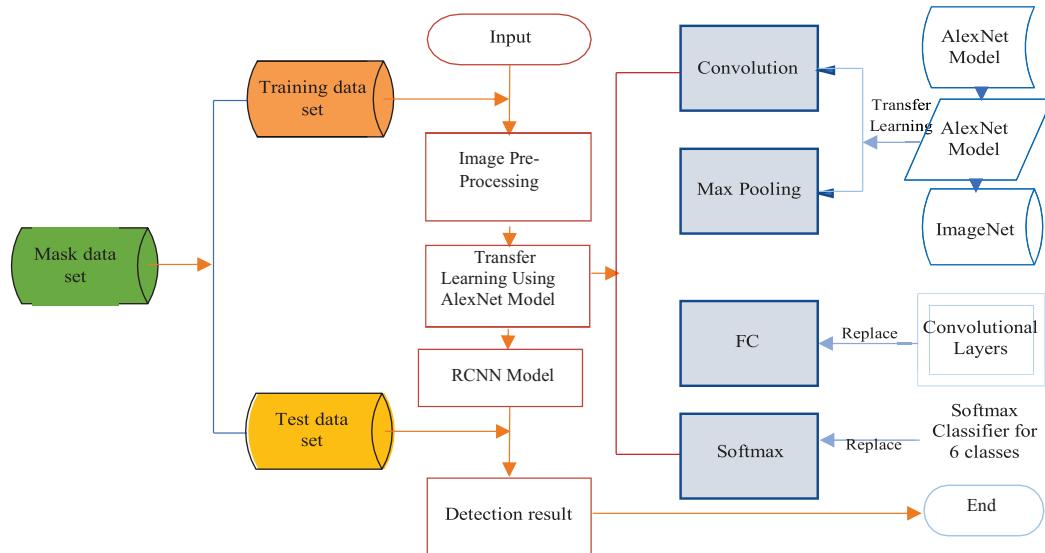


Fig.5. Architecture of RCNN

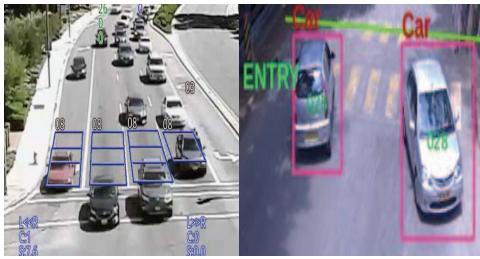


Fig.6. Object detection

a. Single Shot Detection (SSD) MobileNet Architecture

Deep learning based approaches aids to detect the unusual behaviour in frames. However, many previous detection of safety helmet methods includes machine and sensor based method. These existing method faces the issue of sensor failure and scene interference. Deep learning based solution which is intended to overcome the restrictions indicated above for detecting safety helmets. Helmet violation can be detected by using SSD Mobile Net. In SSD Mobile Net, the feature maps are extracted using base model of Mobile Net, then convolutional filters are applied to detect the objects. The Mobile Net is the backbone model, which is a feature map extractor. MobileNet is a lightweight network that uses depth wise separable convolution to reduce the parameters drastically. The fully strided convolutional layer and followed by depth wise strided convolutional layers are used. Batch Normalisation and ReLU are applied to all layers.

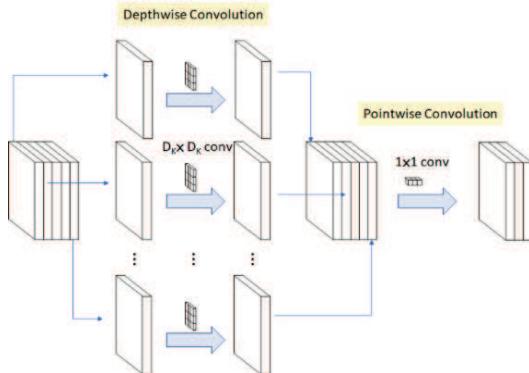


Fig.7. Mobile Net Architecture

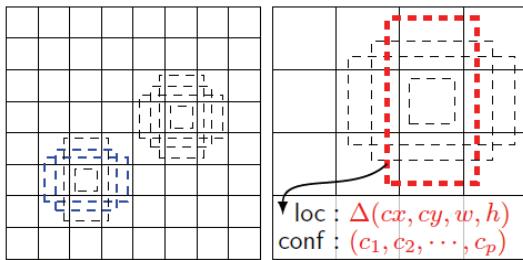


Fig.8. 8*8 feature Map and 4*4 feature Map

The fully connected and SoftMax layer is applied for classification. [17] The depth and pointwise convolution as shown in fig.7. are considered as separate layers in the total number of 28 layers. The final classification in Mobile Net is replaced with SSD head which is made up of several convolutional layers. Several convolution layers are put together on top of the backbone Mobile Net model to provide the bounding boxes over the objects. The SSD detect the boxes with defined sizes in feature map. The detection is performed for different classes with different aspect ratios. There are two types of detections predicted in each default box. The first is, the position of bounding boxes, which includes four offsets and n-1 confidence score, where n object classes. Next one is , the score that indicates the matching of object in that box. A tiny kernel is used to detect the lowest prediction parameter that yields the coordinates of the default box. A non- maximum suppression approach is applied to make the final detection from the set of overlapping detections. The maximum class confidence score will be used to confirm the object classes. Based on the confidence score, the boxes are organised in descending order with the first ones being maintained. Then, the left prediction boxes are filtered using the non-maximum suppression strategy.

One of the difficulties in traffic violation is detecting a motorcycle and helmet in the frame. The shape of the object (motorcycle) in the image, the recognition of persons riding on a motorcycle, locating the biker's head, and the detection of a helmet at the biker's head are challenges one. The image of a biker wearing a helmet and without one is then manually classified from the video data.



Fig.9. Biker with helmet and without helmet

Based on width, height, xmin, ymin, xmax and ymax coordinates of bounding box, a motorbike with a rider and helmet area are annotated and kept under the label called "Biker with helmet" and another label called "Biker with no

"helmet" with an area of a motorcycle with a motor cyclist who does not wear a helmet. 764 images with a size of 227x227x3 are used from Kaggle dataset for training and validating the model.

b. SSD VGG16 Architecture

The VGG16 allows many convolutional layers, Max pooling for down sampling, linear unit (ReLU) for better image feature extraction, fully connected and Soft max layer for classification. VGG16 has 3x3 conv kernels with a stride of one , 2x2 max pool kernels. The base network initially develops parameters for a 1000-category classification [18] as shown in Fig.10. The VGG16 is the backbone model in which final fully connected layer is replaced with SSD head which is placed on the backbone model of VGG16 to detect and locate the objects. The back bone of VGG16 provides extracted feature map to SSD. In SSD , Six convolutional layers performs object detection and classification. SSD will predict 8732 bounding boxes using six convolutional layers. Based on the confidence score, the boxes are organised in descending order with the first ones being maintained. Then, the left prediction boxes are filtered using the non-maximum suppression strategy. Non Max suppression is applied to remove the duplication. The object detection is very difficult in traditional methods due to the different types of mask and the environment condition as well as changes in the environment light. SSD -VGG model could be used for object detection and localization.

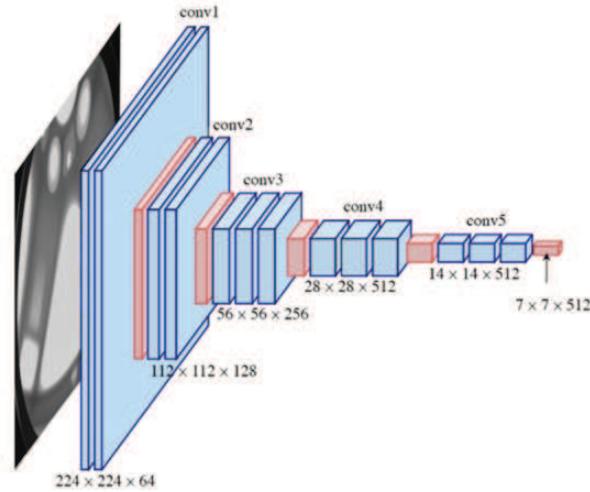


Fig.10. Backbone Model using VGG 16

Additionally, SSD head layers are added to the pre-trained model of VGG-16 network to classify and locate the target. Its parameters are shown in table1. Objects with ground truth boxes are given as input to SSD. Multiple default boxes of different sizes and aspect ratio across the entire image is created for better coverage of location. SSD will predict 8732

bounding boxes for each object in the image. The convolutional layers of SSD check the aspect ratio of each boundary box. SSD match the boundary box which has highest overlap on ground truth box with the help of Intersection Over Union. This Overlap should be greater than 50%. It will pick only maximum overlapping boundary box with ground truth box. Then, Non-Max Suppression is applied. The new SSD Head layer with base layer locates the wearing masks accurately by transferring the learned weightage in base model of VGG16. SSD VGG 16 is used to classify the wearing mask or without mask as shown in fig 11.

Table 1- SSD HEAD

Layer	Grid Size	Kernel Size
Conv 6	19*19	3*3*1024
Conv 7	19*19	1*1*1024
Conv8	10*10	1x1x256 3x3x512-s2
Conv9	5*5	1x1x128 3x3x256-s2
Conv10	3*3	1x1x128 3x3x256-s1
Conv11	1*1	1x1x128 3x3x256-s1



Fig.11. Face Mask Detection

IV. Results and Discussion

The huge amount of data set is essential one to train the deep learning network which will run on GPU processor to improve the testing and training speed. The effect and reliability are used for evaluation of the object detection model. Error rate, precision and recall, testing speed are usually indicators. The algorithm's fast test time is of critical practical importance, because the algorithm takes time for testing a single image. Precision is calculated based on False

Positive (FP) and True Positive (TP), which is defined in equation (1). The number of actual targets that comes under the positive class is measured by precision, whereas recall and error rate is calculated as following equation (2) and (3). The predictions of number of positive classes in out of all positive examples in the dataset is measured by recall. The models which is designed to detect the object are good, since the precision stays high as the value of recall is increased.

$$\text{Precision}(P) = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{Precision}(P)} \quad (2)$$

$$\text{Errorrate} = \frac{\text{FN}}{\text{Precision} - \text{Recall}} \quad (3)$$

a. Mean Average Precision

The AP (Average Precision) metric value is often computed for recall values ranging from 0 to 1, which is used to assess the accuracy of object detectors such as the R-CNN, SSD, and others. The Mean Average Precision is calculated to measure the accuracy of object detector of RCNN, SSD Mobile Net, SSD VGG16 overall classes. The performance of the proposed model is compared with existing models as shown in fig. 12. The proposed model of SSD Mobile Net , SSD VGG16 performs much better at high and low MAP than the other state of the art detectors.

b. Intersection Over Union

IOU (Intersection Over Union) measures the overlap between two borders, which determines the overlapping of anticipated border with the real object boundary. This Overlap should be greater than 50%. IOU measures the correctness of the prediction based on underlying ground truth. It is comparing similarity between anticipated border and the real object boundary $X, Y \subseteq S \in \mathbb{R}^n$ is attained by:

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|} \quad (4)$$

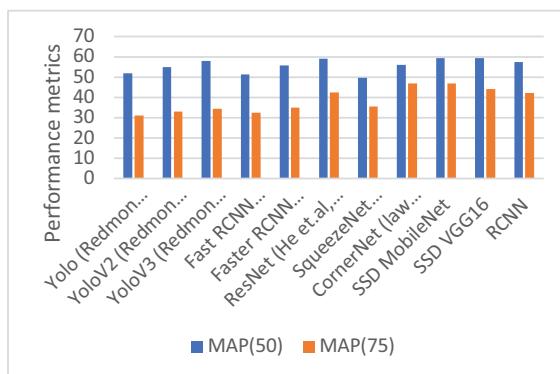


Fig.12. Performance Metrics using MAP

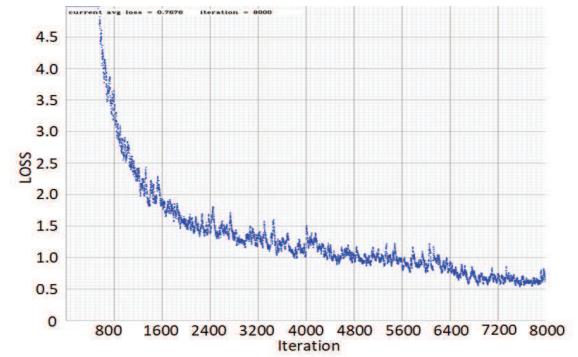


Fig.13. Loss over the iteration

By using tensor board software package, loss metrics are done to properly monitor and analyze the model performance on the dataset. A dropout layer is placed after each convolutional layer, with the value set to 0.25. It deactivates 25% of the hidden layer nodes at random to avoid over-fitting the network. In multi-class classification tasks, target should fit into one of many possible categories. The learning rate may be the most crucial hyperparameter in constructed neural network, which implies the speed at which weights are changed. Here, the learning rate value = 0.003 is used to get better model, that explains the changes of model weights based on predicted error.

V. Conclusion

The RCNN, SSD Mobile Net , SSD VGG16 model's performance are examined for vehicle recognition and classification using Image input. The Transfer learning model provides high accuracy results by employing Alex Net, Mobile Net, VGG16 model for object detection. The accuracy of detection of traffic violation is improved through the transfer learning.

REFERENCES

- [1] Cai, Y., Sun, X., Wang, H., Chen, L., Jiang, H.: Night-time vehicle detection algorithm based on visual saliency and deep learning. *J. Sens.* 2016 (2016)
- [2] Dong,Z.,Pei,M.,He,Y.,Liu,T.,Dong,Y.,Jia,Y.:Vehicle type classification using unsupervised convolutional neural network. In: 2014 22nd International Conference on Pattern Recognition (ICPR), pp. 172–177 (2014)
- [3] Fan, Q., Brown, L., Smith, J.: A closer look at faster R-CNN for vehicle detection. In: 2016 IEEE Intelligent Vehicles Symposium (IV), pp. 124–129 (2016).
- [4] Ferguson, Max & ak, Ronay & Lee, Yung-Tsun & Law, Kincho. (2017). Automatic localization of casting defects with convolutional neural networks. 1726-1735. 10.1109/BigData.2017.8258115
- [5] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size”, 2016

- [6] Feris, R.S., et al.: Large-scale vehicle detection, indexing, and search in urban surveillance videos. *IEEE Trans. Multimed.* 14(1), 28–42 (2012)
- [7] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587 (2014)
- [8] He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(9), 1904–1916, (2015).
- [9] Hei Law, Jia Deng," CornerNet: Detecting Objects as Paired Keypoints", 2018.
- [10] Q. W. Yao. Car face classification based on convolution neural network. Zhejiang University, 2016.
- [11] Kim K., Kim P., Chung Y., and Choi D., "Multi- Scale Detector for Accurate Vehicle Detection in Traffic Surveillance Data," IEEE Access, vol. 7, pp. 78311-78319, 2019. \
- [12] Liu,D.,Wang,Y.:Monza:image classification of vehicle make and model using convolutional neural networks and transfer learning. <http://cs231n.stanford.edu/reports/2015/>
- [13] Rajalakshmi, N.R,Saravanan,K, "Cloud-Based Intelligent Internet of Medical Things Applications for Healthcare Systems", Evolving Role of AI and IOMT in the Healthcare Market, Springer, 2021.
- [14] Rajalakshmi, N.R, ArulKumaran, G, "Virtual Machine Consolidation For Performance And Energy Efficient Cloud Data Center Using Reinforcement Learning", International Journal of Engineering and Advanced Technology, ISSN:2249 – 8958, Volume-8, Issue-3S, Pp. 779-784, February 2019.
- [15] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, pp. 91–99 (2015)
- [16] Su,B.,Shao,J.,Zhou,J.,Zhang,X.,Mei,L.,Hu,C.: The precise vehicle retrieval in traffic surveillance with deep convolutional neural networks. *Int. J. Inf. Electron. Eng.* 6(3), 192(2016).
- [17] <https://bit.ly/33TDhxG>
- [18] Tian, B., et al.: Hierarchical and networked vehicle surveillance in its: a survey. *IEEE Trans. Intell. Transp. Syst.* 16(2), 557–580 (2015)
- [19] Uijlings, J.R., Van De Sande, K.E., Gevers, T., Smeulders, A.W.: Selective search for object recognition. *Int. J. Comput. Vis.* 104(2), 154–171 (2013)
- [20] Wang, H., Cai, Y., Chen, L.: A vehicle detection algorithm based on deep belief network. *Sci. World J.*2014,e647380 (2014)
- [21] Wang K. and Liu M., "Object Recognition at Night Scene Based on DCGAN and Faster R- CNN," IEEE Access, vol. 8, pp. 193168-193182, 2020.
- [22] Zitnick, C.L., Dollár, P.: Edge boxes: locating object proposals from edges. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) *ECCV* 2014. LNCS, vol. 8693, pp. 391–405. Springer, Cham (2014). doi:10.1007/978-3319-10602-1_26..

Performance Evaluation of Convolution Neural Network for Lung Cancer Detection

Dhasn Lydia M

School of Computing

SRM Institute of Science and Technology

Chennai, India

dhasnlydia@gmail.com

Prakash M

School of Computing

SRM Institute of Science and Technology

Chennai, India

prakashm2@srmist.edu.in

Abstract - Cancer is the dangerous diseases nowadays among humans. Particularly lung cancer is the very common one among men and woman. The statistics says, there is a malignancy in men out of 5 and 1 out of 9 in women. Even though it's dangerous, due to the advances in medical field, lung cancer can be treated if it's found at the earlier stage through tomography (CT) screening. However, radiologist felt difficult to detect the malignant lung nodule from CT Image which usually take more time. Thus, the Computer-aided diagnosis systems (CAD) developed to address this issue. But the accuracy and performance of CAD system is still need more improvement. Nowadays deep learning techniques has been use by the researchers to improve the performance lung cancer screening with CT images using CAD system. In this work we reviewed the recent deep learning architectures for lung cancer detection and present the wide-ranging study of standard methods and also the best model has been evaluated using LUNA16 dataset.

Keywords: lung cancer; deep learning; nodule detection; convolutional neural networks; malignant; benign.

I. INTRODUCTION

In 91 of 172 nations, cancer is the primary or succeeding major reason of death beforehand the age of seventy, and it ranks third or fourth in another 22. Lung cancer was the principal reason of death in males in 2018 and 3rd greatest root of death in women, according to the IARC. Prostate and colorectal cancers, as well as cancers of the liver and stomach, are the most frequent tumours in males and the foremost reason of cancer malignance death. In 2018, about 9.6 million cancer fatalities were reported, with 1.8 million deaths attributable to lung malignance cancer reporting for 18.4% of all deaths [1]. Because lung cancer is such a deadly illness, it's essential to catch it early. There is a possibility of increasing survival rate of patient by 50% if it's detected at an initial stage before spreading [2]. On radiological imaging, a lung nodule is a spherical or irregular opacity surrounded by aerated lung that is an important predictor of lung cancer. [3]. Because solitary pulmonary nodules (SPNs) have a high likelihood of becoming malignant nodules, early identification of SPNs is critical for early-stage lung cancer diagnosis [4]. Lung cancer is divided into four phases. The stage 1 cancer is controlled in lungs. The stages II and III cancer not restricted with lungs instead

it will reach and limited to chest. (with bigger and more aggressive tumors categorized as stage III). In final (IV) the Cancer will extent from the chest to supplementary parts of the body.

A lung nodule (or mass) is a tiny abnormal spot that can be discovered during a chest CT scan. These scans are performed for a variety of purposes, including lung cancer screening and checking the lungs if you have symptoms. Most of lung nodules identified on CT scans are not cancerous. They are more commonly caused by previous infections, scar tissue, or other factors. However, testing is frequently required to ensure that a nodule is not cancerous.

Positron Emission Tomography (PET) scan, Chest radiography (x-ray), magnetic resonance imaging (MRI scan), computed tomography (CT), needle biopsy and sputum cytology are some of the procedures used to detect lung cancer [5,6]. According to studies, using LDCT scans to detect those at greater lung cancer risk saved more lives than using chest x-rays. In those who are at a higher risk of lung cancer, getting yearly LDCT scans when symptoms show can help reduce the risk of dying from the disease [6].

As a result, creating CAD systems that might improve radiologist productivity while potentially reducing false-negative findings is a major problem. This computer-aided design technique also aids in the avoidance of potentially harmful histology biopsies. (a) lung segmentation, (b) lung disease quantification, and (c) differential diagnosis are the three processes of a CAD system used to examine lung CT scans. The initial procedure involves identifying the lung boundary, separating the lobes, and, in certain circumstances, detecting and removing the broncho vascular tree. Different tissue abnormalities are found in the second stage. In the last stage, the results are combined to give a likely differential diagnosis. The following process segmentation, feature extraction, and classification algorithms are used to classify lung nodules using image processing method. In this study we focus mainly on the 2nd stage that categorizes lung nodule with distinctive deep learning method especially on CNN. Deep learning enhances the computer's precision and performance nowadays. Using the provided base-truth, the CNN may learn the relationship between characteristics and cancer on its own. Once trained, the network must be in the state

where it able to generalize its knowledge to identify cancerous tumors (or patient-level cancer) in novel conditions that it has not experienced before. The ability to accomplish end-to-end identification by learning the far more significant properties throughout training is a fundamental benefit of deep learning in CAD systems. The current convolutional neural network for lung cancer classification is reviewed in this paper.

Traditional image processing approaches, which include the following procedures, are used in the initial identification of cancer by CAD. In first step the image was pre-processed to locate the particles. In second step standard segmentation technique is used to segment the cancer nodule. The image was then processed to extract properties such as diameter, area, pixel mean intensity and centroid. The classification step follows this differentiate the tumor into benign and malignant one [7].

Recently, neural networks have been rebranded as 'deep learning,' which is a winning Artificial Intelligence technology for essential tasks including speech recognition, picture classification, and creating genuine, readable sentences. The key benefit of utilizing deep learning is that it enhances the computer's precision and performance when detecting and classifying CT images [8]. In this research, we focus on CNN as a deep learning technique. In this paper we review the recent up-to-date CNN methods for lung cancer finding. Convolutional Neural Networks may more accurately identify and categorise lung cancer kinds in a shorter amount of time, which is critical for establishing a patient's appropriate therapy and survival probability.

II. OVERVIEW OF CAD SYSTEM

Computer-Aided-diagnosis has been a promising area of research in recent decades. CADe (computer-aided detection system) and CADx (computer-aided detection system) are the two primary computational systems that have been created to help radiologists (computer-aided diagnosis system). CADe systems use medical imaging to detect lesions, whereas CADx systems use measurements to determine the severity of the lesions. We discussed the CNN based CAD systems in detail. The general structure of the CAD system was shown in Fig 1.

A. Evaluation metrics of CAD system

To compare the diagnosis performance of the index test to the reference test, standard measures are employed. The index test is an AI-based programme that analyses medical data automatically. As a reference test, the biomedical dataset tagging that was done according to the technique is employed. The metrics of the CAD system are shown in Table 1.

Table 1: Evaluation of matrix of CAD system

Indicator	Formula
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$
False positive rate	1-specificity
Specificity	$TN / (TN + FP)$
Positive Predictive Value	$TP / (TP + FP)$
Negative Predictive Value	$TN / (TN + FN)$
Likelihood ratio of negative result	(1-sensitivity) / specificity
Sensitivity	$TP / (TP + FN)$
Likelihood ratio of positive result	sensitivity / (1-specificity)

III. OVERVIEW OF CONVOLUTIONAL NEURAL NETWORK

Convolutional neural networks, or CNNs, are a form of artificial neural network. CNN is kind of deep learning model for grid patterns such as images that uses numerous building blocks like as convolution, pooling, and fully connected layers to learn feature hierarchies through backpropogation. The first two-layer convolution layer and max pooling layer perform the feature mining and the following layer a completely linked layer maps the extracted feature into the classifier. The comparison of these layers are shown in Table 2.

A. Blocks of CNN

Figure 1 depicts the CNN architecture, which contains blocks of convolution layer, max pooling layer, and one or more fully linked layers.

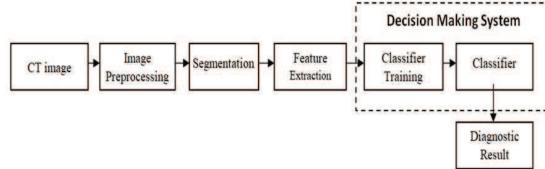


Figure 1: Architecture of CNN
Table 2: Comparison of three layers in CNN

Layers	I/O
Convolutional Layer	3D cube of the previous feature map as input 3D cube with 1 2D map/ filter as output
Pooling Layer	3D cube of aforementioned feature or property map as input 3D cube, 1 map/ filter, condensed spatial dimension as output
Fully Connected Layer	Input: a 3D cube that has been flattened, as well as a prior set of feature maps. 3D cube with one 2D map per filter as output

a. Convolution Layer

A CNN first layer is convolution layer. The convolution layer are the layers where filters are applied to the original images. Convolution layer is usually a combination of linear and nonlinear operation that perform the feature extraction operation [20].

Convolution seems to be a linear method that multiplies a collection of weights by the input. An array of data input and a weighted array termed as both a kernel or filter are multiplied. The result of this multiplication is a single value called a feature map [21]. The result of the linear operation is passed through a nonlinear activation function. This function is mostly used to boost the output's non-linearity. ReLu is a common activation function.

a. Pooling Layer

A pooling layer designed to address the problem of down sampling. This implies that even tiny changes in the location of feature in the picture outcome in a novel feature map.

After the convolution layer, a pooling layer is added, which can be repeated one or more times throughout the model. To construct a new pooled feature map, a pooling map is functional to every set of the feature map. Pooling is the process of choosing a pooling technique that shrinks the volume of the feature map.

a. Fully Connected Layer

The last layer of a network is the completely linked layer. A feed forward neural network is another name for this type of system. The output of the pooling or convolution layer is the fully linked layer's input. A fully linked layer's output normally has the same number of output nodes as the number of classes.

b. Activation Function

A node at the end or in the centre of a neural network is called the activation function. Sigmoid, tanh, ReLu, maxout, and ELU are examples of distinct types of activation functions. ReLu is the most often used activation function in neural networks. The fact that ReLu activation does not stimulate all neurons at the same time is the most important advantage. ReLu activation is quicker than sigmoid and tanh activation [22].

IV. DIFFERENT ARCHITECTURES OF CNN

All CNN architecture follows the same general design by successively applying the convolutional layer to the input, periodically down sampling the spatial dimension. The classic network architecture and modern network are general forms of networks. The classic networks are LeNet-5, AlexNet, VGG 16 and the modern networks architectures are Inception, Xception, ResNet, ResNeXt, DenseNet.

LeNet-5

It is the simplest architecture has 2 convolutional and 3 fully connected layers . The architecture has about 60,000 parameters.

AlexNet

It has 8 layers – 5 convolutional layer and 3 fully connected layer. This was the first model to implement the Rectified Linear Units (ReLu). This architecture has 60M parameters.

VGG-16

13 convolutional layers, 3 fully linked layers, and a ReLu unit are included in this model. This design requires 500MB of storage space and contains 138M parameters.

Inception-v1

This architecture consists of 22 layers with a total of 5 million parameters. This architecture is build using modules instead of stack of convolutional layers. Later version Inception V3 uses the batch normalization.

ResNet-50

This architecture has 26M parameters and it is one of the early adoption of batch normalization. The building blocks of Res Net includes Conv and identity blocks. The Table 3 enlist the study of various convolutional neural network used to diagnose lung cancer.

V. EVALUATION OF CNN ARCHITECTURE

Two methods proposed by Anum Masood et al [37] and Cheng-Jian Lin et al [38] has given good accuracy in classifying the lung nodules in this study. These two methods has been evaluated using LUNA16 dataset to check its trustworthiness. Both the methods nearly reached the accuracy as shown in table 4 which is claimed by the authors. The model of the above method given in equation 1,2,3.

$$\zeta_c(w, h | \phi) = \sum_{(a,b) \in bin(w,h)} z_{w,h,c}(a + a_0, b + b_0 | \phi) / n \quad (1)$$

where ϕ denotes parameters of the network, $\zeta_c(w,h | \phi)$ is the relevance score of $(w,h)^{th}$ bin to malignant category c, z , w,h,c is the score map generated by last convolutional layer, (a_0,b_0) is the top-left corner of ROI, and n denotes the total pixel number in the bin.

$$\zeta_c(\phi) = \sum_{w,h} \zeta_c(w, h | \phi) \quad (2)$$

$$\xi_c(\phi) = \exp(\zeta_c(\phi)) / \sum_{i=0}^5 \exp(\zeta_i(\phi)) \quad (3)$$

Figure 2 shows the accuracy of two standard methods with LUNA16 dataset which is almost near to the claimed accuracy by the authors.

Table 3: Studies on state-of-the-art CNN architecture for lung cancer detection

Year	Reference	Benchmark	CNN Architecture	Results	Keypoints
2015	Wei Shen et al [23]	LIDC-IDRI	MCNN	Accuracy-86.84%	MCNN capture nodule heterogeneity from stacked layers. class specific features are learned by concatenating response neuron activation. SVM and RF classifiers are used for classification
2016	Marios Anthimopoulos et al [24]	ILD from Univ Hospital of Geneva, Bern Univ Hospital	CNN	Accuracy-85.6%	5 layer convolution is used with LeakyReLU activations, followed by average pooling
2016	Prajwal Rao et al [25]	LIDC	CanNet	Accuracy-76.0%	CanNet based on the layers of the CNN along with the appropriate value for network parameters, followed by pooling layer, dropout layer and fully connected layer
2016	Wei Li et al [26]	LIDC-IDRI	CNN	sensitivity-87.1%	convolution neural network have two convolutional layer, downsampling layer after each layer and fully connected layer is appended to the last downsampling layer. The ROI region is recognised as nodule or nonnodule by the output probabilities.
2017	Wafaa Alakwaa et al [27]	Kaggle Data Science Bowl, LUNA16	Modified U-Net,3D-CNN	Accuracy-86.6%, Misclassification-13.4%, FP-1.9%, FN-14.7%.	a modified U-Net trained on LUNA16 data was used to first detect nodule candidates in the Kaggle CT scans, then 3D -CNN to classify the CT scan as positive or negative for lung cancer.
2017	Qi Dou et al [28]	LUNA16	3D CNN	Sensitivity-94.4%	Employed 3D-CNN for false positive reduction
2018	Wentao Zhu et al [29]	LIDC-IDRI	3D Faster R-CNN and GB machine	Accuracy-81.41%	3D Faster R-CNN is designed for nodule detection and gradient boosting machine is used for nodule classification.
2018	Fang [30]	LIDC-IDRI	GoogleNet	Accuracy-81%, sensitivity-84%	GoogleNet with transfer learning approach was built and median intensity projection was used for multi-view feature
2018	Haritha Sathyam [31]	LIDC-IDRI	AlexNet	Accuracy-98%	Three classes of lung images are classified using this AlexNet
2018	LIU Lu et al [32]	LIDC-IDRI	CNN,particle swarm optimization optimized SVM	Accuracy-91.94%	CNN has been used to extract the features of the lung nodule, these feature dimension are reduced by principle component analysis method and classification of nodule features are done by particle swarm optimization SVM.
2018	Silva et al [33]	LIDC-IDRI	CNN based on PSO	Specificity 98.2%, Accuracy-97.62%, Sensitivity 92.2%	Network performance was enhanced using particle swarm optimization that optimise the network parameters
2019	Haichao Cao et al [34]	LUNA16	MBEL - 3D - CNN	CPM score is 87.3%	Multi-branch ensemble model includes three models namely VeggNer,IResNet and DenseNet. multi learned network architecture combines the learning feature to detect lung nodule.
2019	Nasrullah et al [35]	LIDC-IDRI,LUNA16	CMixNet,R-CNN	sensitivity-94%, Specificity-91%,FROC Score-94.21%	The identification and categorization of lung nodules is done using a customized mixed link network. The nodule detection was done using RCNN and further classification was performed by Gradient Boosting Machine (GBM)
2019	Qianqian Zhang et al [36]	LICD-IDRI	VGG16,VGG19, Xception, ResNet50, MobileNet, NASNetMobile, DenseNet121, and NASNetLarge	Highest Accuracy :87.7%, Sensitivity 82.73%, Specificity 92.38%	This perform a end-to-end classification from raw 3D nodule CT patches.State-of-the-art CNN models are modified to 3D-CNN model.experimental result shows that DenseNet and Xception produce better result for lung nodule classification.

Year	Reference	Benchmark	CNN Architecture	Results	Keypoints
2020	Anum Masood et al [37]	SPH6, LIDC-IDRI, ANODE09, LUNA16	3DDCNN and mRPN	sensitivity-98.4%, specificity-92%, AUROC-96%, accuracy-98.5% with 2.1 FPs per scan	Multi-region proposal network is used for automatic selection of region of interest. 3D Deep convolution neural network enthused by ResNet-101 is used for reducing false positive rate. VGG16 is the basic layer of this 3DDCNN. The score map generated by the last layer is used to find the malignancy score. This CAD system is improved by integrating cloud computing provided by virtual machine and software as a service.
2020	Cheng-Jian Lin et al [38]	LIDC-NCI	Taguchi-based CNN	Accuracy-99.6%	Taguchi method used the orthogonal table method for selecting preliminary factor. This method is combined with AlexNet to improve the accuracy of original AlexNet.
2021	Amrita Naik et al [40]	LUNA	FractalNet	Accuracy-94.7%, Specificity-90.41%, Sensitivity-96.68%	FractalNet is a CNN that coordinate the sub paths of varying length and use filters to transform signal before send it to the subsequent layers.

Table 4 : Accuracy with LUNA16 Dataset

Model	Accuracy	Difference
3DDCNN and mRPN	98.45 %	0.05
Taguchi-based CNN	99.5 %	0.1

The difference between claimed accuracy and the calculated accuracy with LUNA16 dataset of two methods are 0.05 and 0.1 respectively which is shown in figure 3.

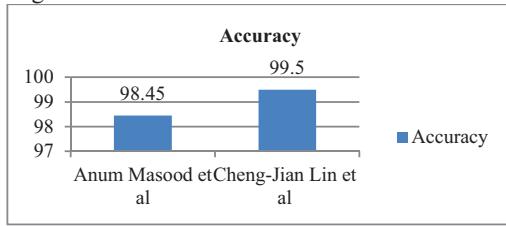


Figure 2: Accuracy of Two standard methods with LUNA16 DB

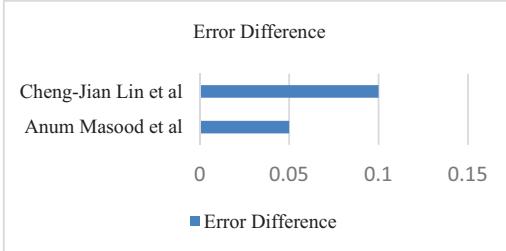


Figure 3: Error difference of Two standard methods with LUNA16 DB

VI. CONCLUSION AND FUTURE WORK

This paper has been written to help the researchers to understand the CAD system and general evaluation metrics usually used in analyzing the CAD system performance. Deep learning-based architecture

particularly CNN architectures in lung cancer detection have been studied critically and their results also compared. This paper also focuses on evaluating some standard CNN architectures with different datasets to check the said accuracy of the authors. The result also shows that, the claimed accuracy is truthful and that can be achieved when it's implemented in the system.

Even though these methods give the good accuracy using used datasets, evaluating with different datasets and real data's is essential to validate the claimed accuracy by the authors. This would help to improve the trustworthiness of these methods and possibly used in the development of CAD systems. Collecting some real data of CT images for evaluating the above methods and proposing the better model which would consistently give standard accuracy for any dataset is much needed and that will be taken as a future work.

References

- [1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., & Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries: global Cancer Statistics 2018. CA: a Cancer Journal for Clinicians, 68. <https://doi: 10.3322/caac.21492>
- [2] Song, Q., Zhao, L., Luo, X., Dou, X., 2017. Using deep learning for classification of lung nodules on computed tomography images. Journal of Healthcare Engineering 2017. <https://doi.org/10.1155/2017/8314740>
- [3] Hansell, D.M., Bankier, A.A., MacMahon, H., McLoud, T.C., Muller, N.L., Remy, J., 2008. Fleischner society: glossary of terms for thoracic imaging. Radiology 246, 697-722. <https://doi: 10.1148/radiol.2462070712>.
- [4] H. T. Winer-Muram, "The solitary pulmonary nodule," Radiology, vol. 239, no. 1, pp. 34-49, 2006. doi: 10.1148/radiol.2391050343.
- [5] Kyaw, N.N., Naing, K.K., Thwe, P.M., Htay, K.K. and Htun, H., Detection and Classification of Lung Cancer Stages using Image Processing Techniques. International Journal of Electrical Electronics & Computer Science Engineering, Volume 6, Issue 4 (August, 2019).

- [6] Lung Cancer Early Detection, Diagnosis, and Staging, cancer.org | 1.800.227.2345. American Cancer Society.
- [7] Jeyaprakash Vasanth Wason & Ayyappan Nagarajan., Image processing techniques for analyzing CT scan images towards the early detection of lung cancer. Bioinformation. 2019; 15(8): 596–599. doi: [10.6026/97320630015596](https://doi.org/10.6026/97320630015596).
- [8] QingZeng Song, Lei Zhao, XingKe Luo, and XueChen Dou., Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. Hindawi Journal of Healthcare Engineering Volume 2017, Article ID 8314740, <https://doi.org/10.1155/2017/8314740>.
- [9] Jun Gao, Qian Jiang, Bo Zhou and Daozheng Chen., Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. AIMS Mathematical Bioscience and Engineering. 16(6): 6536–6561. DOI: <https://10.3934/mbe.2019326>.
- [10] H. Fujita and D. Cimr, Computer Aided detection for fibrillations and flutters using deep convolutional neural network, Inf. Sci., 486 (2019), 231–239.
- [11] W. Li, Y. Zhao, X. Chen, et al., Detecting Alzheimer's Disease on Small Dataset: A Knowledge Transfer Perspective, IEEE J. Biomed. Health. Inf., 23 (2019), 1234–1242.
- [12] L. Martin, J. Jendeberg, P. Thunberg, et al., Computer aided detection of ureteral stones in thin slice computed tomography volumes using Convolutional Neural Networks, Comput. Biol. Med., 97 (2018), 153–160.
- [13] H. C. Shin, H. R. Roth, M. Gao, et al., Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, IEEE Trans. Med. Imaging, 35 (2016), 1285–1298.
- [14] O. F. Ahmad, A. Soares, E. B. Mazomenos, et al., Artificial intelligence and computer-aided diagnosis in colonoscopy: Current evidence and future directions, Lancet. Gastroenterol. Hepatol., 41 (2019), 71–80.
- [15] González-Díaz, DermaKNet: Incorporating the Knowledge of Dermatologists to Convolutional Neural Networks for Skin Lesion Diagnosis, IEEE J. Biomed. Health. Inf., 23 (2019), 547–559.
- [16] U. Raghavendra, H. Fujita, S. V. Bhandary, et al., Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images, Inf. Sci., 441 (2018), 41–49.
- [17] M. Anthimopoulos, S. Christodoulidis, L. Ebner, et al., Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network, IEEE Trans. Med. Imaging, 35 (2016), 1207–1216.
- [18] L. C. S. Afonso, G. H. Rosa, C. R. Pereira, et al., A recurrence plot-based approach for Parkinson's disease identification, Future Gener. Comput. Syst., 94 (2019), 282–292.
- [19] H. Pratt, F. Coenen, D. M. Broadbent, et al., Convolutional Neural Networks for Diabetic Retinopathy, Procedia. Comput. Sci., 90 (2016), 200–205.
- [20] Yamashita, R., Nishio, M., Do, R. and Togashi, K., 2021. *Convolutional neural networks: an overview and application in radiology*.
- [21] Brownlee, J., 2021. *How Do Convolutional Layers Work in Deep Learning Neural Networks?*.
- [22] Medium. 2021. *Basic Overview of Convolutional Neural Network (CNN)*.
- [23] Shen, W., Zhou, M., Yang, F., Yang, C. and Tian, J. (2015). Multi-scale Convolutional Neural Networks for Lung Nodule Classification. *Lecture Notes in Computer Science*, [online] pp.588–599.
- [24] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe and S. Mougiakakou, "Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1207-1216, May 2016, <https://doi.org/10.1109/TMI.2016.2535865>.
- [25] P. Rao, N. A. Pereira and R. Srinivasan, "Convolutional neural networks for lung cancer screening in computed tomography(CT) scans," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016, pp.489-493, <https://doi.org/10.1109/IC3I.2016.7918014>.
- [26] Wei Li, Peng Cao, Dazhe Zhao, Junbo Wang, "Pulmonary Nodule Classification with Deep Convolutional Neural Networks on Computed Tomography Images", Computational and Mathematical Methods in Medicine, vol. 2016, ArticleID 6215085, 7 pages, 2016. <https://doi.org/10.1155/2016/6215085>.
- [27] Alakwaa, W., Nassee, M. and Badr, A., 2021. Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN), Vol. 8, No. 8, 2017.
- [28] Q. Dou, H. Chen, L. Yu, J. Qin and P. Heng, "Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1558–1567, July 2017, <https://doi.org/10.1109/TBME.2016.2613502>.
- [29] W. Zhu, C. Liu, W. Fan and X. Xie, "DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification," 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 673–681, https://doi.org/10.1109/WACV_2018.00079.
- [30] T. Fang, "A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections," 2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET), 2018, pp. 286–290, <https://doi.org/10.1109/CCET.2018.8542189>.
- [31] H. Sathyan and J. V. Panicker, "Lung Nodule Classification Using Deep ConvNets on CT Images," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1–5, <https://doi.org/10.1109/ICCCNT.2018.8494084>.
- [32] Liu Lu, Liu Yapeng, and Zhao Hongyuan. 2018. Benign and Malignant Solitary Pulmonary Nodules Classification Based on CNN and SVM. In *Proceedings of the International Conference on Machine Vision and Applications (ICMVA 2018)*. Association for Computing Machinery, New York, NY, USA, 46–50. DOI:<https://doi.org/10.1145/3220511.3220513>
- [33] Giovanni Lucca Franc a da Silva, Thales Levi Azevedo Valente,Aristofanes Corr ea Silva, Anselmo Cardoso de Paiva, Marcelo Gattass .2018. Convolutional neural network-based PSO for lung nodule false positive reduction on CT images. *Computer Methods and Programs in Biomedicine* (2018), <https://doi.org/10.1016/j.cmpb.2018.05.006>.
- [34] H. Cao *et al.*, "Multi-Branch Ensemble Learning Architecture Based on 3D CNN for False Positive Reduction in Lung Nodule Detection," in *IEEE Access*, vol. 7, pp. 67380–67391, 2019, <https://doi.org/10.1109/ACCESS.2019.2906116>.
- [35] Nasrullah, N.; Sang, J.; Alam, M.S.; Mateen, M.; Cai, B.; Hu, H. Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies. *Sensors* 2019, 19, 3722. <https://doi.org/10.3390/s19173722>.
- [36] Qianqian Zhang, Haifeng Wang, SangWon Yoon, Daehan Won, Krishnaswami Srihari. 2019. Lung Nodule Diagnosis on 3D Computed Tomography Images Using Deep Convolutional Neural Networks. *International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing* August 9–14, 2019, 363–370.
- [37] Masood *et al.*, "Automated Decision Support System for Lung Cancer Detection and Classification via Enhanced RFCN With Multilayer Fusion RPN," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7791–7801, Dec. 2020.
- [38] Lin, C.-J.; Jeng, S.-Y.; Chen, M.-K. Using 2D CNN with Taguchi Parametric Optimization for Lung Cancer Recognition from CT Images. *Appl. Sci.* 2020, 10, 2591.
- [39] Bin Sun, Fengyin Liu, Yusun Zhou, Shaolei Jin, Qiang Li, and Xinyu Jin. 2020. Classification of Lung Nodules Based on GAN and 3D CNN. In *Proceedings of the 4th International Conference on Computer Science and Application Engineering (CSAE 2020)*. Association for Computing Machinery, New York, NY, USA, Article 112, 1–5.
- [40] Naik, A., Edla, D.R. & Kuppili, V. Lung Nodule Classification on Computed Tomography Images Using Fractalnet. *Wireless Pers Commun* **119**, 1209–1229(2021).

Decision Trees to Detect Malware in a Cloud Computing Environment

1 st VIJAYARAJ Associate Professor, Department of Computer Science and Engineering <i>Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India</i> vijaiphdraj@gmail.com	2 nd SUMATHI M Assistant Professor, Department of School of Computing <i>SASTRA Deemed to be University Thanjavur, Tamilnadu</i> sumathi@it.sastra.edu	3 rd RAJKAMAL M Application Developer IBM, Bangalore, <i>India</i> rajkamalmurugasean@gma il.com	4 th UGANYA VIJAYARAJ uganyaee@gmail. com Assistant Professor, ECE, Saveetha School of Engineering, <i>SIMATS,</i>	5 th DT.Kamaleshwar Assistant Professor, kamalesh4u2@gmail.com Department of Computer Science and Engineering <i>Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai,</i>	6 th D.Rajalakshmi Assistant Professor rajisacet@gmail.com Department of CSE Sri Sairam Institute of Technology, Chennai
---	---	---	---	--	--

Abstract— Cloud computing is the most demanding technology for efficient storing of information and provide facilities online. As the data is stored in huge amount, it is very much prone to cyber-related attacks. Using this increasing technology to protect storage information which is computer based from malware attacks and show up several rewards along with traditional detection schemes. Because cloud-based protected products are portable, they may be used on any computer-based system, including Cyber Physical Systems (CPS), personal devices such as Internet of Things (IoT) devices, laptop, desktop and mobile devices. The malicious software is computer software that is designed to launch exploitable assaults against computer systems that gains unauthorized access to the system which confidentiality, integrity and authentication (CIA) requirement triad agreement has been distributed, then the whole organization will be in trouble is generally called as malware. To detect various forms of malware attacks, the effective malware detection (MD) technique is proposed in the work in a cloud based data storage environment. The proposed MD technique will be learnt with the dataset that contains characteristics on different virtual machines (VM) which identify distinctive MD features in an efficient manner. The decision tree (DT) is then given selected features. For identifying malware and benign samples, a DT employs learning and rule-based (LRB) agents. The proposed technique can identify any software with high detection rate and accuracy. It will be characterized based on the features, and falls on two buckets – malware and benign.

Keywords— Malwares, Decision Tree, Benign, Internet of Things, Cyber Physical System.

I. INTRODUCTION

I. INTRODUCTION

Due to digital technology growth cyber-attacks are massively increase in recent years. Various virus varieties are the most typical cause of cyber-attacks. In cyber-attacks, malware is a type of computer programme designed to take advantage of a computer's weaknesses and execute nefarious acts in order to get financial gain. Malware comes in many forms, including infection, worm, Trojan, backdoor, root package, and emancipate product. For various purposes, the harmful malwares are created by the process of code variant [1]. Based on the nature of malware code variants different types of actions are performed like to steal sensitive

information, launch distributed denial of service attacks and allow inaccessible code execution. The majority of the more recent assaults take data from people, such as credit card points of interest on account management frameworks, jumble processing data on various devices/drives to fragment victims' process the framework, and hurt more than millions of users around the world [2].

MD is a technique for determining if a programme is malicious or not. Heuristic, behavior, and model checking are examples of traditional methodologies, whereas cloud computing, deep learning techniques, and mobile device-based MD's are examples of emerging approaches. The advantages of a cloud-based MD methodology over conventional techniques are numerous. The cloud-based environment offers easy access, on-demand capacity, high computational control, and significantly larger databases offer low costs. Using different VM's and host servers, different execution trails of the same virus have been obtained [3]. Individual PCs, flexible devices, and IoT devices benefit from the cloud environment, which enhances detection rates. Different discovery calculations will be implemented in different type servers in the future. Limited algorithms enhances detection performance while lowering the number of false positives and negatives not all type of detection algorithms. Machine learning algorithm-based malware detection approaches are now widely used for effective detection. Machine learning (ML) approaches, when compared to traditional methods, produce better outcomes due to the feature selection process. For MD, feature selection algorithms select the high impact attributes from the dataset [11].

ML algorithms are using in MD are, support vector machine (SVM), random forest (RF), DT and Naïve Bayes (NB) etc. The rate of accuracy in ML algorithm depends on the feature selection algorithm and training model. Additionally, the ML algorithms in MD is to handle the unknown malwares effectively, automate the malware sample extraction, efficient identification of unknown malware variations, and reduce human effort and time requirement [12]. Different types of MD techniques are behavior-based detection, signature-based detection, heuristic-based detection, mobile -based detection, IoT-based detection, model checking-based detection, and ML-based

detection. Among these detection, ML based detection techniques produces better classification result. Due to these benefits, the ML based ML is proposed in this work.

The overall organization of this article as: Section II, focused on literature review of the existing malware detection technique with the merits and demerits. In section III, the proposed decision tree based malware detection technique discussed with necessary architecture and algorithm. In section IV, the proposed technique experimental results are compared to the existing MD techniques and section V, the conclusion and future enhancement is discussed.

II. RELATED WORKS

L.Xiao, Y.Li, and colleagues proposed a cloud-based MD approach for mobile devices that uses off-loading. This strategy works well on mobile devices but is ineffective on larger devices such as desktop PCs and laptops. In cloud contexts, Babu and Murali devised and constructed an intermediary malware defence system. This method reduces computing time and costs. The work's weakness was that traffic delays increased while the model was being trained on the data [4].

Shen et al. combined cloud and fog computing with a MD system implemented by an intrusion detection system (IDS). This method reduces data traffic delay and data transmission overhead. However, this strategy does not appear to be effective against all types of computer viruses [10]. For malware detection, Zhou and Yu et al. used a cloud-assisted model, as well as a dynamic approach to prevent malware spread. This strategy effectively and visibly prevents the transmission of harmful code. However, this strategy only worked in a limited number of VMS [5].

The malware detection approach in cloud infrastructure was proposed by Abdelsalam et al. It is also possible to detect system calls, paths and resources, and various file types. This method is incapable of detecting traditional viruses [7]. Mirza et al. suggested an intelligent malware detection technique based on a cloud-based energy-efficient hosting strategy. This method outperforms traditional anti-virus software. However, for larger data, the computational time is significant. Deynannis et al suggested trustAV, a cloud-based MD method. Even on remote networks, this MD method protects transfer and processing of user data. There is a lot of data to work with, and computing takes time [6].

Gupta S et al. proposed the future smart connected communities to fight COVID-19 outbreak. It helps a lot after covid era with lot of new technologies being added. Its not efficient on the conventional virus. N.A.Azeez et al. used windows PE MD using collaborative learning. This technique works well efficiently in personal computers and desktop related windows operating system version [8]. O.Aslan et al. discussed a review of cloud-based MD system opportunities, advances and challenges. This technique provided solutions for many new challenges which are very helpful in further advancements. These advancements yet to be implemented and that requires a lot of data and time. Based on the literature review the following problems are identified in the exisiting techniques [9].

The major challenges of ML based MD process are:

- High computational cost requirement – train the classifier in a frequent interval to detect the new malwares.
- Adversarial ML - The malware developer try to penetrate the malware detector.

Existing technique Problems:

A DT is a decision-making technique which uses a tree-based structure to organize prediction options and the corresponding potential outcomes, like chance event outcomes, asset costs, and utility. In this approach to demonstrate an algorithm that has conditional control statements built in. A DT is a flow-graph like structure. In DT each internal nodes represent a "test" on a quality (like coin flip comes up heads or tails) and each branch represents the outcome of the test, finally each leaf node represents a class name (shows the choice taken after computing all traits). In DT the paths that lead from the root to the leaf are classified according to certain rules. The choice trees (DTs) are a non-parametric directed learning strategy utilized in regression and classification. The major objective is, to construct a model for predicting the esteem of a target variable by learning straightforward DT rules deduced through the information features. DT gives more productivity and efficiency than the existing framework.

III. PROPOSED DECISION TREE BASED MALWARE DETECTION

The user sends an executable file in a computer network for storing in cloud storage. The provided file is then run on a virtual machine, and the results are collected using dynamic tools, and the characteristics are extracted and mapped against the trained characteristics. Based on the results the network renders to results page after submission and user get to know whether the file is malware or legitimate. If it is benign it is uploaded into cloud i.e AWS boot S3 and vice-versa. The flow diagram and algorithm of the proposed work is shown in figure 1 and Algorithm 1. The malware dynamic features (MDF) and list of predefined properties (LPP) are taken as an input.

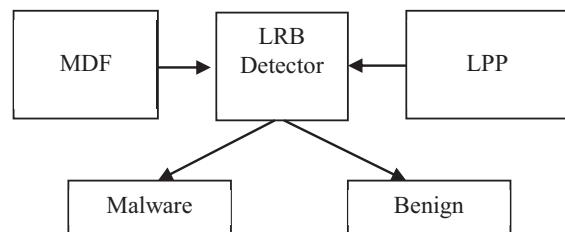


Figure 1. Flow Diagram of Proposed Work

Algorithm : Malware Detection Using Decision Tree

Input: MDF, LPP

Output: Identified malware samples

Procedure:

1. Path_score{‘low’, ‘medium’, ‘high’}
2. Frequency $\leftarrow \{\text{few}, \text{avg}, \text{high}, \text{excessive}\}$
3. If path_score = low and frequency == high or excessive then

4. SuspiciousFile = Malware
 5. Else
 6. SuspiciousFile = Benign
-

The LRB based detection algorithm takes LPP and MDF as an input and produces the marked program samples as output. Based on this output, we set the entropy value and based on this the important samples will be extracted and those will be sent to the model for training. So, based on the samples it got trained, it will be able to say whether the file given to it as an input is a malicious or not. Algorithm judges the output based on the path score and frequency category of the dataset whether the path score is low or moderate or high and whether the frequency category is few or average or many or excessive.

In the LRB detection technique, the rules are frame from the characteristics of malwares. By using regular expressions and features, the powerful and complex detection rules are framed. The sample rules are listed as follows:

Rule Malware {

metadata:

Description = "This is malware"

Strings:

\$A1= "GetProcAddress <CreateThread>
LdrGetProcedureAddress"

\$A2= "Getsystemmetrics"

\$A3= "FindresourceExw LoadResource"

\$A4= "NtOpenFile"

Condition:

Any of them

}

A benign signature generated using the following rule.

Rule Benign {

metadata:

Description = "This is benign"

Strings:

\$A1= "www.micorsoft.com"

\$A2= "23.211.9.92"

\$A3= "216.58.205.164"

\$A4= "www.google.com"

Condition:

Any of them

}

By using this rule based identification technique, the malware and benign are easily identified. The experimental results are discussed in the upcoming section.

Merits of the Proposed Technique:

- Cloud environment is very secure when compared to the existing traditional file storage system which is not vulnerable to cyber-attacks.
- Cloud environment provides more computational power especially in case of MD and even storage at lower costs.
- The system don't have to think about hardware resources in cloud environment. Everything will be taken care by the CSP.

IV. EXPERIMENTAL RESULTS

In a proposed work, the frontend of the webpage is designed by the elements like HTML, CSS and JS. In this page, the file uploading option is given to upload a dataset. By using flask app, the home page url is entered. In the initial stage, the dataset from the virus share website where the dataset file contains various MD5 hashes and behaviors. After that, the important features are extracted from the dataset and stored in the list which will be used later for classification. From 54 features, 13 features are considered as important and extracted from the feature set. These features are extracted by the rule based classification and tree classifier.

```
C:\Users\innoc\Desktop\Project>python learning.py
Researching Important feature based on 54 total features

13 features identified as important:
1. feature DLLCharacteristics (0.166201)
2. feature Machine (0.109128)
3. feature Characteristics (0.001832)
4. feature Subsystem (0.070802)
5. feature MajorSubsystemVersion (0.065994)
6. feature SectionsSubEntropy (0.059237)
7. feature ImageBase (0.051329)
8. feature ResourcesMaxEntropy (0.049269)
9. feature VersionInformationSize (0.0466842)
10. feature SizeOfOptionalHeader (0.035414)
11. feature ResourcesMinEntropy (0.033454)
12. feature MajorOperatingSystemVersion (0.027152)
13. feature StackReserve (0.026655)

Now testing algorithm
DecisionTree : 99.127128 %

DecisionTree with a 99.127128 % success
Saving algorithm and feature list in classifier directory...
Saved
False positive rate : 0.626911 %
False negative rate : 1.444217 %

C:\Users\innoc\Desktop\Project>
```

Figure 2. Feature Selection

DT are simple decision making diagrams, what the input is and what the corresponding output is in the training data. The major objective of the DT is to develop a training model which can be used to predict the class or value of the target variable by learning simple decision rules inferred from prior training data. Python language is used for implementation of this work. By using pandas and numpy functions, the data analysis and numerical operations are performed. The 80:20 ratio used as training data ratio and testing data ratio process. Figure 2 shows the MDF selection process of the proposed work. Table 1 shows the list of extracted features from the dataset.

TABLE I - Feature Identification

Sl.No	Feature Name	Feature Score
1	DllCharacteristics	0.166201

Sl.No	Feature Name	Feature Score
2	Machine	0.09120
3	Characteristics	0.081832
4	Subsystem	0.070025
5	MajorSubSystemVersion	0.065994
6	SectionMaxEntropy	0.059237
7	ImageBase	0.051329
8	ResourceMaxEntropy	0.049269
9	VersionInformationSize	0.0046042
10	SizeofOptionalHeader	0.035414
11	ResourceMinEntropy	0.033454
12	MajorOperatingSystemVersion	0.027152
13	SizeofStackReserve	0.026655

From this analysis there are 13 features identified as important and extracted from the dataset. These extracted features are stored as a pickle file and further loaded to second phase such as training phase. Figure 3 shows the benign file. If the file is not an injected file or it is an actual legitimate file, those files are classified as benign file and accepted for uploading process.

Performance measures are used to measure the proposed system efficiencies are

- True Positive (TP) – TP specify the ratio of correctly identified benign as benign.
- True Negative (TN) – TN represents the ratio of malware samples are identified as malware.
- False Positive (FP) - FP represents the ratio of malware classified as benign.
- False Negative (FN) – FN represents the ratio of benign classified as malware.
- Accuracy – represents the correctly identified ratio of malware and benign to the total samples.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN})$$

- Precision – represents the ratio of correctly identified positive samples and the total identified positive samples.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

- Recall – represents the correctly identified positive samples and the total identified samples of the actual data.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

- F1 Measure – represents the average value of the precision and recall.

$$\text{F1 Measure} = 2 * ((\text{Precision} * \text{Recall}) / (\text{Precision} * \text{Recall}))$$



Figure 3. Benign File

Figure 4 shows that if the file is an injected file or malicious file, the corresponding file updation will be rejected in file uploading process. The classification value of the proposed technique shown in table 2.

TABLE II – CLASSIFICATION VALUES

Class Type	Precision	Recall	F1-Measure	Support
Class_0	0.968	0.978	0.975	90187
Class_1	0.956	0.907	0.942	482149

Figure 2 shows the detection accuracy comparison of the proposed and other classification techniques. When compared to other techniques, the proposed technique prediction accuracy is high. Hence, the proposed technique classify the malware in an efficient manner is proven.

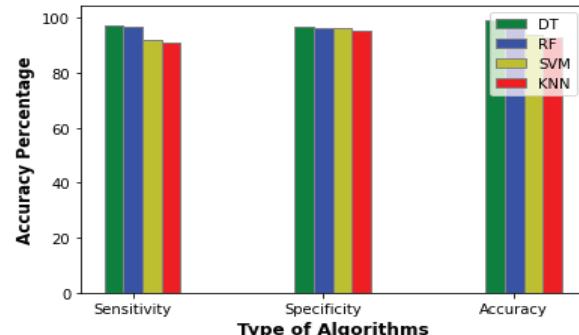


Figure 2. Malware Detection Comparison

V. CONCLUSION AND FUTURE ENHANCEMENTS

This proposed work presents a MD solution for cloud computing consists of two parts such as client and cloud environment. From the client side the files are uploads to the cloud and reviews the findings, which reveal whether the file samples are malicious or benign. Getting dataset from virus share and handling them is a complex task. In MD, the model detecting whether a host system contains malicious file or not. If the system fails to detect the malware, our system is very prone to cyber-attacks which the whole data will be in risk. Hence, it is necessary to have a malware detection system in every system. Here, the cloud environment as a database which is very secure and easily accessible instead of storing on our system. Cloud is also very cheap and also will be having lot of specifications which can be easily accessible. The data can also be trained

with other classification algorithms like RF and SVM for getting better accuracy. To extend this project into mobile or web application where it detects suspicious files and warns the user.

REFERENCES

- [1] Q. K. A. Mirza, I. Awan, and M. Younas, "CloudIntell: An intelligent malware detection system," *Future Gener. Comput. Syst.*, vol. 86, pp. 1042–1053, Sep. 2018.
- [2] N. M. Babu and G. Murali, "Malware detection for multi cloud servers using intermediate monitoring server," in Proc. Int. Conf. Energy, Commun., Data Analytics Soft Comput. (ICECDS), Aug. 2017, pp. 3609–3612.
- [3] D. Deyannis, E. Papadogiannaki, G. Kalivianakis, G. Vasilidis, and S. Ioannidis, "TrustAV: Practical and privacy preserving malware analysis in the cloud," in Proc. 10th ACM Conf. Data Appl. Secur. Privacy, Mar. 2020, pp. 39–48.
- [4] W. Zhou and B. Yu, "A cloud-assisted malware detection and suppression framework for wireless multimedia system in IoT based on dynamic differential game," *China Commun.*, vol. 15, no. 2, pp. 209–223, Feb. 2018.
- [5] S. Shen, L. Huang, H. Zhou, S. Yu, E. Fan, and Q. Cao, "Multistage signaling game-based optimal detection strategies for suppressing malware diffusion in fog-cloud-based IoT networks," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1043–1054, Apr. 2018.
- [6] M. Abdelsalam, R. Krishnan, Y. Huang, and R. Sandhu, "Malware detection in cloud infrastructures using convolutional neural networks," in Proc. IEEE 11th Int. Conf. Cloud Comput. (CLOUD), Jul. 2018, pp. 162–169.
- [7] M. Yousefifi-Azar, L. G. C. Hamey, V. Varadharajan, and S. Chen, "Malytics: A malware detection scheme," *IEEE Access*, vol. 6, pp. 49418–49431, 2018.
- [8] Das Malwerk. *Malware Downloading Website*. Accessed: Jan. 15, 2021. [Online]. Available: <https://dasmalwerk.eu/>
- [9] J. Jeon, J. H. Park, and Y.-S. Jeong, "Dynamic analysis for IoT malware detection with convolution neural network model," *IEEE Access*, vol. 8, pp. 96899–96911, Jun. 2020.
- [10] Tekdefense. *Malware Downloading Website*. Accessed: Jan. 15, 2021.[Online]. Available: <http://www.tekdefense.com/downloads/>
- [11] Kamalakanta Sethi, Rahul Kumar, Lingaraj Sethi, Padmalochan Bera, Prashanta kumar patna, "A novel machine learning based malware detection and classification framework", 2019 International conference on cyber security and protection of digital services, 2019, PP 1-4.
- [12] Jagsir Singh, Jaswinder Singh, "A survey on machine learning based malware detection in executable files", Journal of Systems Architecture, 2021, PP 1-21.

EVALUATING PERFORMANCE METRICS IN CLASSIFYING BITCOIN MIXING SERVICES USING DECISION TREE ALGORITHM

¹Ezhilmathi S

Ph.D Research Scholar

Department of Computational Intelligence
School of Computing
SRM Institute of Science and Technology
Kattankulathur, Tamilnadu-603203
es5083@srmist.edu.in.

²Selvakumara Samy S

Assistant Professor

Department of Computational Intelligence
School of Computing
SRM Institute of Science and Technology
Kattankulathur, Tamilnadu-603203
selvakus1@srmist.edu.in

Abstract—Bitcoin is a decentralized peer-to-peer (P2P) cryptocurrency system with an innovative payment network. For blockchain applications, ECDSA is formed with public and private keys, especially in Bitcoin, which uses an elliptic curve in a cryptography standard known as Secp256k1 to ensure funds are spent by legitimate owners. Despite the fact that ECDSA is a key component of Bitcoin transactions, today's criminals employ pseudonymous addresses, which make it impossible to track unlawful activity because they don't maintain real-world identities. Bitcoin Fog, Helix Mix, and other mixing services are supposed to provide transaction privacy. However, these services are commonly utilized for concealment, it is more difficult to track criminals. First, using time-frequency analysis, features can be recovered at the network, account, or transaction level to create attributed temporal heterogeneous network Motifs and characterize many forms of address patterns. To increase the performance of evaluation metrics, we introduced a decision-tree machine learning approach for the categorization of bitcoin mixing services from unlabeled addresses in this work. Experiments with bitcoin Kaggle datasets are frequently used to determine the success of our categorization approach.

Keywords— Bitcoin, ECDSA, Blockchain, elliptic curve, Mixing services

I. INTRODUCTION

Blockchain is a system that combines decentralized data storage, shared transmission, agreement instruments, cryptographic calculations, and other advances to provide decentralization, information permanence, and programmability. Since Satoshi Nakamoto [16] first introduced Bitcoin's concept in 2008, the blockchain technology has been widely used in industries such as advanced money, smart manufacturing, storage network the board, anti-counterfeiting information management, and so on. At the same time, blockchain security has attracted a lot of attention from the academic community.

Bitcoin is a digital currency that allows semi-anonymous participants in inaugurations to participate without the need for a centralized middleman. Bitcoin was founded using blockchain technology [1], and it is the result of a long period of increased research and advancements in fundamental improvements and computations [2]. Bitcoin functions as an interpersonal association in a variety of ways, but one that is distinguished by the anonymity of the impersonators and the plutocrat-related capabilities rather than the traditional social communication of its constituents. Bitcoin has garnered critical scrutiny because to its content, with studies such as [3] delving into exemplifications of Bitcoin value shift and [4] examining a large number of association packages since the initial three times of Bitcoin trade data. Each user in the Bitcoin organization can work with various Bitcoin addresses, which can assist to improve the security of

Bitcoin users. One trial was conducted to link Bitcoin locations to the personalities of the owners in order to assess its security. One study reenacted the Bitcoin framework in a meeting setting, with the possibility to differentiate 40 of the attendees [5]. Another study linked Bitcoin users to similar resources such as SatoshiDICE and Wikileaks [6]. The first stage towards honoring the owners of unlabeled Bitcoin addresses is to group the labels depending on ownership; that is, the Bitcoin labels are grouped with other Bitcoin tendencies claimed by a related individual or organization. Whenever this occurs, the ID of anyone in a given group will aid in the connection of different labels in that group to the true personality.

Bitcoin is a cryptocurrency that is not controlled by a central authority. Mining exchange records in an attempt to illustrate their exchange designs can help identify mixing facilities. Because of inadequate spot data, dynamic cycles with different exchanges, and different confusion designs, the Bitcoin mixing facilities [7] are extremely unsafe. Mixing facilities are distinguished from the web level, information level, and exchange level using an element-based organizational framework. ATHN themes are used to identify bitcoin mixing. PU learning issues are the current model for bitcoin mixing recognition [8], which works with reserve following and wrongdoing location in the Bitcoin environment and achieves a high TPR and low FPR. They made half-and-half themes, remembering the need of fleeting themes for AAIN and ATHN themes in TAIN for mixed discoveries.

The goal of this research is to assess the performance metrics for unlabeled address classification in Bitcoin mixing services such as Bitcoin fog, Helix Mix, and others. The following is the work's major provocation:

1. Using the Elliptic Curve Digital Signature Algorithm(ECDSA), generate a public and private key for signing and verifying the addresses.
2. We used Attributed Temporal Hybrid Network Motifs (ATHN Motifs) to extract the features utilizing timeline and frequency series analysis.
3. Implement a decision tree system to evaluate the performance indicators and classify Bitcoin mixing services like Bitcoin Fog and Helix Mix from unlabeled addresses.

Analyze the transaction record returned by the wallet explorer for generalization issues and to inject noise into datasets. Creating several nodes for a transaction-address interaction network, which is one of the most powerful techniques for network mining, Addresses can be categorized as labelled or unlabeled to classify them in mixed services. By using mixing services, one can send money to more than 5 different addresses without having to

go through KYC. This can be solved using one of the machine learning methods, such as the decision tree approach, which can significantly enhance performance in three parameters: F1-score, precision, and recall.

The rest of the research work illustrates what to look out for. The work on the decision tree algorithm is presented in Section II. Section III-V provides an overview of loading datasets and classifying addresses before proposing a feature extraction approach based on a detection model. Section VI brings this effort to a close.

Using kaggle.com, the mixing services from bitcoin transactions were extracted. By establishing attributed temporal hybrid network motifs where each node can have a transaction or an address that contains both homogeneous and heterogeneous network nodes, the transaction data has a number of addresses with labelled and unlabeled addresses. The general framework of measuring performance measures in Bitcoin Mixing Services is depicted in Figure 1. One of the most essential techniques for the interaction of network nodes is network motifs.

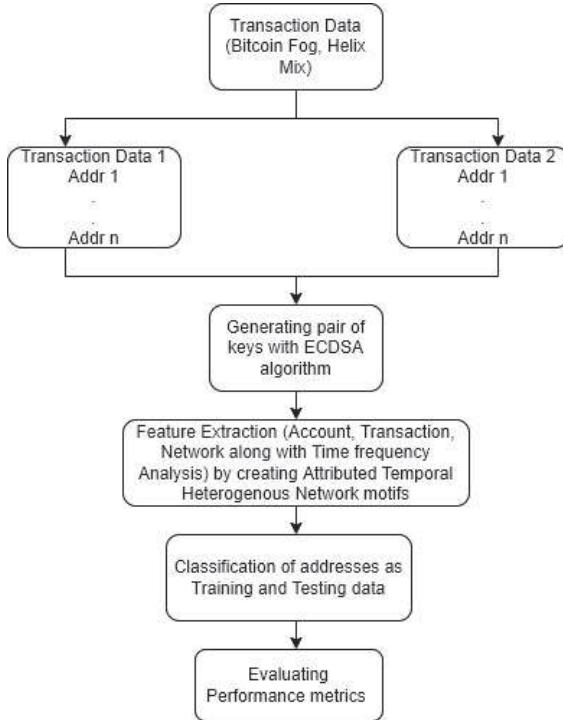


Fig.1 Overall structure of evaluating Performance metrics in Bitcoin Mixing services

II. RELATED WORKS

Through multifariousness and analysis of over 6000 reports, Casagrandeet.al [9] fashion was applied on a spectacular frame contextual disquisition. We've seen that his technique produces space models of comparable quality to physically constructed area models in a more intelligent manner. Learning frame with centers to work out that cyberbullying may be severe direct that attempts to control or naughtiness others misleading is a teary illustration of destructive lead that mercilessness should do as well as have the possibility for a reappraisal, according to Ezhilmathi S et.al., [13].

With an emphasis on protection and framed security objects, Enas Al Kawasmie et al. [14] provide a fabric-based design model for a Decentralized Carbon Emigrations Trading Structure (D-CETI). The concept and execution are carried out in response to the challenge of anonymously exchanging fossil fuel byproducts among trading agents. Despite the use of aliases, Androulaki E et al., [15] presents a number of security concerns because all transactions are disclosed openly in the framework. Ezhilmathi S et al. [17] propose a clever part-extract-based solicitation approach for the unfathomable universe of Twitter jokes. Appraisal-related components, accentuation-related elements, syntactic and semantic elements, and model characteristics are among the four social affairs elements that have been removed.

Zengyou He et al. [18] propose a new-tree grounded sequence bracket method for constructing a succinct decision tree from point space that includes all subsequences in the training sequences. This will improve scalability and delicacy with various decision trees, as well as investigate the possibility of constructing arbitrary wood.

Tie Qiu et al. [19] expand similarity indicators to directed weighted networks and extract a series of similarity indicators between bumps and edges in this paper. Then, utilizing decision tree ensembles, a directed edge weight vaticinator model with increased vaticinator delicacy and robustness was developed. Exploring additional significant aspects in the heterogeneous network to improve the delicacy of prognostications.

The difficulty of providing interpretable and correct data for categorization was stated by ZYashesh Dhebar et al. [20]. Without using any synthetic data, the non-linear decision tree has reached exceptional delicacy. Including non-polynomial and generic terms in split-rule expressions.

III. LOADING DATASET

For Bitcoin block explorers, datasets may be collected from kaggle.com by assigning names to addresses, making exchanges with a few administrations, and watching how the Bitcoin streams consolidate. However, the Kaggle data name data set has not been updated since roughly 2016, implying that Kaggle does not include new arising administrations. On kaggle.com, the Bitcoin organization has a large number of exchanges with bitcoin transaction details that are publicly available. For the year 2019, we have Bitcoin exchange data. There are around 150, 000 exchange records in each preview. The dataset was created using Kaggle label information. The loaded dataset could be further filtered to separate Bitcoin mixing services like Bitcoin Fog and Helix Mix from unlabeled Bitcoin addresses.

The table I below provides Bitcoin transaction details for the year 2019. The data is loaded into the machine for classification of mixing services like Bitcoin Fog and Helix Mix, among others, after it is retrieved from the Kaggle. Some of the mixing services listed below are unable to trace the details of a bitcoin transaction.

TABLE I Bitcoin Transaction detail for the year 2019

Date	Total bitcoins	Bitcoin Txn	Bitcoin output volume
11/10/2019 0:00	1339450	26958	3550
11/11/2019 0:00	1342900	27036	93450
12/1/2019 0:00	1424550	28674	3300
12/2/2019 0:00	1429450	28773	14950
1/1/2019 0:00	1629350	32806	5950
1/2/2019 0:00	1635850	32938	6600
2/1/2019 0:00	1881950	37919	12300
2/2/2019 0:00	1889050	38061	7100
3/1/2019 0:00	2162150	43663	10300
3/2/2019 0:00	2171950	43867	12400
4/1/2019 0:00	2421500	49054	10924.2
4/2/2019 0:00	2428900	49209	8800
5/1/2019 0:00	2700300	58695	41344.6
5/2/2019 0:00	2708800	58892	12403
6/1/2019 0:00	2946100	64894	14142.01
6/2/2019 0:00	2952750	65076	16201.97
7/1/2019 0:00	3186200	71649	47384.78
7/2/2019 0:00	3194950	71888	39069.81
8/1/2019 0:00	3580250	98174	275826.17
8/2/2019 0:00	3592950	98587	80756.88
9/1/2019 0:00	3879000	110213	25273.89
9/2/2019 0:00	3886550	110521	37692.37
10/1/2019 0:00	4155850	123547	74309.81812
10/2/2019 0:00	4161300	124048	46316.14777
11/1/2019 0:00	4451600	137743	61627.41
11/2/2019 0:00	4460000	138026	65171.14
12/1/2019 0:00	4746900	201953	84893.23
12/2/2019 0:00	4756550	201925	80412.8

1. Bitcoin fog

Bitcoin fog, which was created in 2011, is one of the longest-running plutocrat laundering installations on darknets. It can be tough to figure out who owns a bitcoin address. Each pullout can be split down into an arbitrary number of sales for a given period of time using this setup.

2. Helix Mix

Helix Mix is a 2017 mixing service that provides two sorts of mixing services: regular and light. Customers were typically asked to construct a wallet, after which the bitcoins sent to the wallet were mixed and sent to random addresses. The light versions of Bitcoin can be transferred to up to five different addresses.

IV. MATERIALS AND METHODS

1. Generating Pair of Keys

The Elliptic Curve Digital Signature Algorithm is a type of asymmetric cryptography utilized by Bitcoin's blockchain (ECDSA). It's a cryptographic and elliptic curve math-based safe digital signature system. To understand how Bitcoin works, we need to grasp a few crucial features concerning ECDSA.

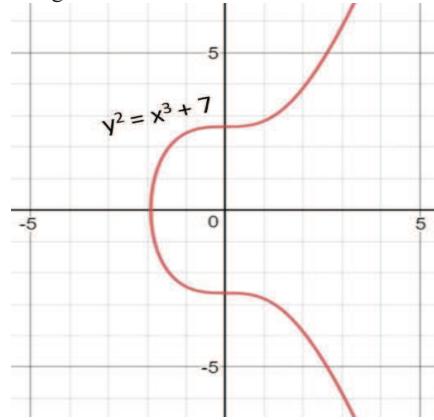


Fig 2 Elliptic curve

The conventional geometric representation of an elliptic curve is shown in Figure 2 [12]. It's not an exact representation of an elliptic curve, but it's essential for scalability. Bitcoin, for example, employs a specific elliptic curve and a set of mathematical constants described in the Secp256k1 cryptography standard. The National Institute of Standards and Technology produced this cryptography standard (basically the governing body of cryptography).

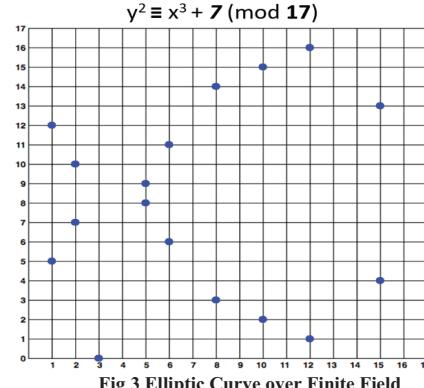


Fig 3 Elliptic Curve over Finite Field

Figure 3 depicts an elliptic curve over finite fields. The Secp256k1 curve is used by Bitcoin to meet the formula $y^2 \equiv x^3 + 7 \pmod{p}$ over \mathbb{F}_p , with $p = 2^{256} - 2^{32} - 2^9 - 2^4 - 1$. This formula yields an elliptic curve with a highly precise shape. It specifies the curve and possible spots for the private and public keys of Bitcoin.

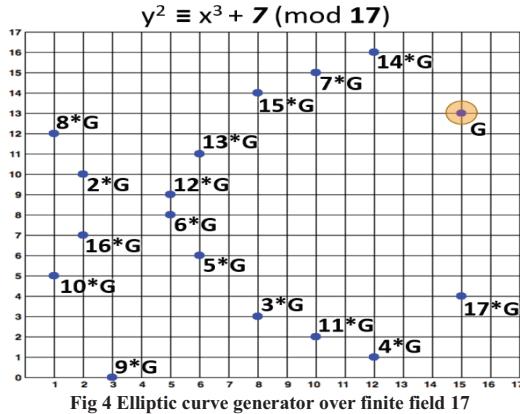
Bitcoin employs the ECDSA algorithm [11], which consists of four steps:

a. Parameter generation

Over the elliptic curve, the process [22] uses an elliptic curve C and a prime order O group generator G. It also uses the hash function H. Finally, several users can access the generator parameters (C, G, O, and H).

b. Key Generation

The process generates a public key $pu = pr * G$ by selecting a private key pr from 2 to O-1 at random.



The amount of times the generator point was multiplied to get the public key is the private key. The Elliptic curve generator is shown in Figure 4 over a finite field. The following is the elliptic curve trap-door function:

Trap-door function:

$$pu = pr * G$$

where,

G = Generator point {x,y}

pr = Private key

pu = Resulting Public key Co-ordinates {x,y}

The private key is used as a variable in the trap-door function for the public key.

2. Temporal Heterogenous Network motifs construction

Organization themes are sub-diagrams that recycle the same thing inside a single organization or across multiple companies. Each of these sub-diagrams, which is defined by a specific example of between vertices relationships, could represent a system in which various capacities are carried out effectively.

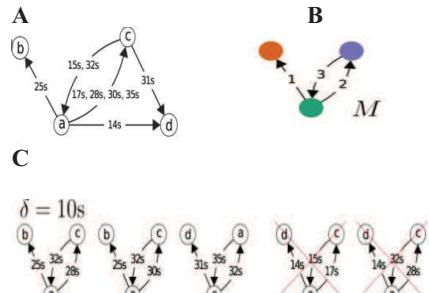


Fig.5 Attributed Temporal Heterogenous Network Motifs

Nine temporal edges in a temporal graph [21]. Each edge's timestamp is stated in seconds. B: 3-node, 3-edge-temporal motif M, for example. The edge labels are in charge of edge ordering. C: For each instance, the temporal motif M is formed with $\delta = 10$ seconds. Due to crossed-out

patterns that are not instances of M, the edge does not appear within or all edge sequences are out-of-order. Figure 5 depicts temporal heterogeneous network motifs that have been assigned.

Temporal network motifs are a sequence of timestamped edges that adjust to a predetermined design, exactly like the predetermined length of time in which the events should occur. The predefined design is actually a coordinated multigraph, and its edges are requested.

3. Feature Extraction with Time-frequency Analysis

a. Network features

Network motifs in an association are small subgraph designs that are more advanced than those in randomised associations [10]. To represent cooperative designs and disclose practical parcels in the network, we recommend employing certain advanced request network highlights (i.e., network motifs). The network vs frequency timeline for bitcoin transactions in the 2019 sample is shown on Fig 6.

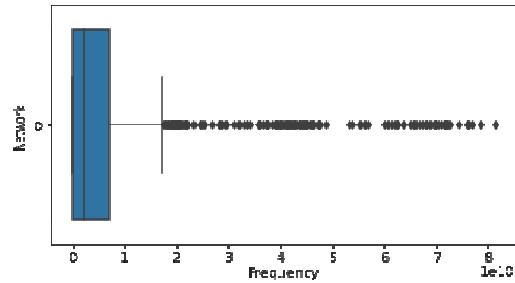


Fig 6 Network Timeline

Both AAIN (account-account interaction network) and TAIN (Transaction-account interaction network), which are homogeneous and heterogeneous in nature, are used to extract the characteristics. Network motifs are minor subgraph patterns in a network that occur more frequently in a randomised network and are referred to as "network motifs."

$$Z-score = \frac{\eta_r - \mu_r}{std(\mu_r)}$$

Where η_r is the frequency, μ_r and $std(\mu_r)$ are the mean and variance related to the Z-score pattern important for network motif.

b. Account features

We create account highlights to show the current state and vibrancy of a location. For example, addresses associated with Bitcoin exchanges typically have a greater exchange recurrence for a large number of businesses, whereas many traditional clients' exchange recurrence is frequently much smaller. The Address versus Frequency timeline for the bitcoin transactions in the 2019 sample is depicted in Figure 7.

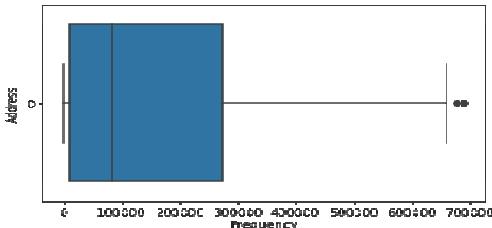


Fig 7 Bitcoin address Timeline

a. Transaction features

Exchange highlights are used to estimate label exchange practises in mixing services. As a result, if the mixing administration simply sends a suspect equal add up to its beneficiaries in the following squares, the relationship between senders and beneficiaries of a mixing interaction would be easily recognised. Many tend to behave as delegate addresses (e.g., hubs) to participate throughout the period spent asset parting and coordinating by mixing administrations. Bitcoins via exchange origin are supplied to the correct beneficiaries after being disseminated by a slew of middle-person addresses over an indefinite period of time. The transaction vs frequency timeline for bitcoin transactions in the 2019 dataset is depicted in Figure 8.

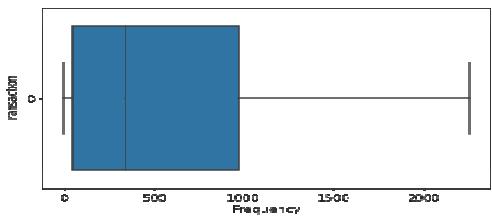


Fig 8 Bitcoin Transaction Timeline

$N = (T, B)$ is a transaction network having a tip T and a border B . The labelled network is denoted by $NL = (T, B, S, L)$, where S is the set of features extracted from N and T is the label set. Each of the $|T|$ samples in $\{S = s_1, s_2, s_3, \dots, s_{|T|}\}$ is described by F features or characteristics written as $s_i = \{s_{i1}, s_{i2}, \dots, s_{in}\}$. $\{L = l_1, l_2, \dots, l_{|T|}\}$ is a label set that provides sample tags $|T|$ for various categories, $l_i \in \{1, 2, 3, 4\}$

b. Time-frequency analysis

To narrate the transaction behaviour of labels, transaction features are extracted. The x-axis in Fig.9 represents the timeline, while the y-axis represents the frequency of bitcoin transactions with address, market value, and the total number of transactions.

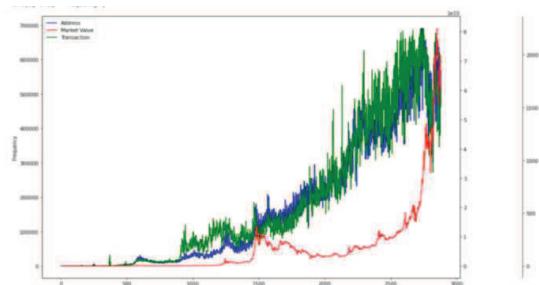


Fig.9 Timeline vs Frequency for Bitcoin transaction with respect to address, market value, Transactions

V. EVALUATION MODEL

We use the widely used decision tree as a classifier (DT). The danger of an operator being mistaken for a positive instance is larger than the risk of being mistaken for a negative instance since covert operative events are virtually always positive. As a result, we can choose an edge based on the asset events' cast probability, and solid negative drawings are chosen from the unlabelled examples if the risk of misinterpretation as a positive instance is low.

Decision trees

Each node represents a feature, each node link represents a decision, and each leaf node represents a class label. A decision tree is a classification strategy for data based on its attributes. It's also used in data mining applications because it's good at processing large amounts of data. Because decision trees do not require domain knowledge, they are perfect for exploratory knowledge discovery.

A decision tree is a type of learning model that is frequently used in ordering. We divide the dataset into at least two sets in this system. The decision tree's inside hubs reflect a test on the highlights, a branch depicts the result, and leaves are decisions made after the result has been handled.

The following is how the Decision Tree fills in:

- As the root of a tree, find the best component of the dataset.
- Divide the data into two sets: training and testing. Subsets should be created data with a comparable element trait in each subset.

Figure 10 depicts the bitcoin dataset's decision tree method generation. It explains how to classify mixing services from unlabeled addresses by using Attribute Selection Measure (ASM) to split the records into smaller subsets and break the dataset into smaller subsets to increase assessment metrics like Accuracy, F1 score, Precision, and Recall. By dividing the data into training and testing sets, the detection model can be evaluated.

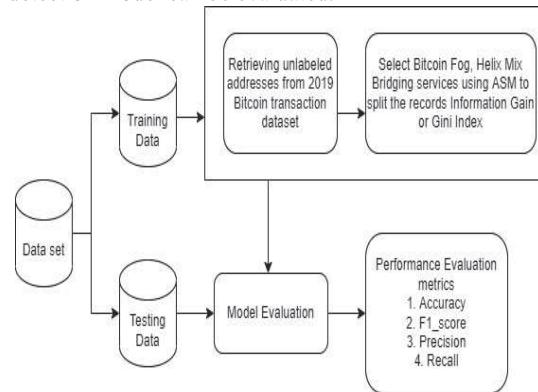


Fig.10 Evaluating the performance metrics

To evaluate the quality of the classifiers, we used the following formulae to calculate the accuracy value, precision, recall, and F measure:

$$\begin{aligned}
\text{Accuracy} &= \text{TP} + \text{TN} / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \\
\text{Precision} &= \text{TP} / (\text{TP} + \text{FP}) \\
\text{Recall} &= \text{TP} / (\text{TP} + \text{FN}) \\
\text{F-Score} &= 2 \times (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})
\end{aligned}$$

Training Set: To get some solid negative occurrences, we frame the training set with 70% unlabeled addresses.

Testing set: The testing set is framed by the remaining 30% of credible negative cases.

The Attribute Selection measure is a splitting rule that aids in determining feature breakpoints. The entropy of a dataset is a measure of its irregularity. Information gain is the polar opposite of entropy, in which actions result in a decrease in entropy. For given trait values, information gain processes the difference between entropy before the split and normal entropy after the split of the dataset. The table II below compares the outcomes of various algorithms that have previously produced them.

Method	TP	FP	TN	FN
DT	6.789	0.897	0.946	1.12
LR [7]	2.987	1.543	3.524	2.42
SVM [8]	2.16	2.312	3.648	2.54

Table II Evaluation of classification with different Methods

The evaluation metrics of bitcoin transaction addresses with respect to DT, LR, and SVM algorithm accuracy, precision, recall, and F1-score have been calculated using the given table values.

Information gain

$$Info(D) = \sum_{i=1}^m \pi \log_2 \pi$$

Where π is the probability of an arbitrary feature in Dataset D of class C and m is number of transactions.

Gini Index

Gini Index gives larger partitioning which uses squared proportion of classes. It works as “ $1 - (P(C1)^2 + P(C2)^2 + \dots + P(Cn)^2)$ ”. It is given by the formula

$$Gini = 1 - \sum_{i=1}^c \pi^2$$

Where c is the class from dataset and π is the probability of arbitrary feature. Because of how they work with noisy or missing data and may be ensembled to construct more powerful predictions, decision trees are one of the most often used ML models. Another benefit of splitting the decision tree is that it allows us to get more exact results.

RESULT AND DISCUSSION

The experimental result demonstrates the efficacy of our suggested model, which uses the ECDSA algorithm to generate a pair of keys. With Timeline frequency analysis, the characteristics may also be retrieved by attribution of temporal heterogeneous network motifs (ATH motifs) such as network features, account features, and transaction features. The decision tree algorithm aids in the

classification of mixing services such as Bitcoin Fog and Helix Mix from unlabeled addresses, as well as improving the confusion matrix accuracy score by up to 82 percent (see Table III). The other metrics (F1 score, Precision, and Recall) all return favorable results of 87%, 88%, and 85%, respectively. The performance of the evaluation metrics by the decision tree method is shown in Figure 11.

Method	Accuracy	Precision	Recall	F1-Score
DT	0.8286	0.883	0.858	0.870
LR [7]	0.6323	0.659	0.552	0.601
SVM [8]	0.5502	0.483	0.460	0.471

Table III Parameter comparison of DT with LR and SVM

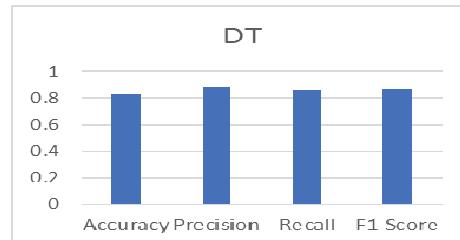


Fig.11 Decision tree Evaluation metrics

CONCLUSION

In this paper, we focus on key generation pairs such as public and private keys, where Bitcoin generates possible points in its public and private keys using the Elliptic Curve Cryptography standard Secp256k1. By producing attributed temporal heterogeneous network Motifs, attributed temporal heterogeneous network Motifs may be extracted at the network, account, and transaction levels using time-frequency analysis to characterize the various sorts of address patterns. The proposed decision tree machine learning approach for classifying bitcoin mixing services from unlabeled addresses has been developed in order to increase the performance of evaluation metrics. Experiments with bitcoin Kaggle datasets are frequently used to determine the success of our categorization approach.

In the future, the Random Forest algorithm might be used to classify mixing services like Bitcoin Fog and Helix Mix from unlabeled addresses in order to improve assessment metrics and compare factors like accuracy, precision, recall, and F1-score with decision trees for the same dataset.

REFERENCES

- [1] A. Murko and S. L. R. Vrhovec (2019) “Bitcoin adoption: Scams and anonymity May not matter but trust into bitcoin security does” in *Proc.3rd Central Eur. Cybersecurity Conf. (CECC)*, New York, NY, USA, pp. 1–6.
- [2] N. Z. Aitzhan and D. Svetinovic (2018) “Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams,” *IEEE Trans. Depend. Secure Comput.*, vol. 15, no. 5, pp. 840–852.
- [3] Y. Yuan and F.-Y. Wang (2018) “Blockchain and cryptocurrencies: Model, techniques, and applications,” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 9, pp. 1421–1428.

- [4] W. Chen and Z. Zheng (2018) "Blockchain data analysis: A review of status, trends and challenges," *J. Comput. Res. Devices*, vol. 55, no. 4, pp. 1853–1870.
- [5] A. Narayanan and J. Clark (2017) "Bitcoin's academic pedigree" *Commun. ACM*, vol. 60, no. 12.
- [6] J. Scott (2017) "Social Network Analysis" 4th ed. Thousand Oaks, CA, USA: SAGE.
- [7] Jiajing Wu et.al., (2021) "Detecting Fraternize Services via Mining Bitcoin Transaction Network With Hybrid Motifs" *IEEE Transactions on Man, Systems & Cybernetics*. DOI:10.1109/TSMC.2021.3049278.
- [8] Tao-Hung Chang and DavorSvetinovic (2020) "Improving Bitcoin Ownership Identification Using Transaction Patterns Analysis" *IEEE Transactions on Man, Systems & Cybernetics: Systems*, Volume: 50, Issue: 1.
- [9] E. Casagrande et.al(2017) "Semiautomatic system domain data analysis:A smart grid feasibility case study" *IEEE Transactions on Systems, Man and Cybernetics*, vol. 47, no. 12, pp. 3117–3127.
- [10] R. Milo et.al.,(2002) "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827.
- [11] <https://eprint.iacr.org/2018/660.pdf>
- [12] <https://wizardforcel.gitbooks.io/practical-cryptography-for-developers-book/content/asymmetric-key-ciphers/elliptic-curve-cryptography-ecc.html>
- [13] Ezhilmathi S et.al., (2021) "Multiclass Words Analysis Using Supervised Learning Technique in Community Network" *Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4*, Pages. 3302 – 3311.
- [14] E. Al Kawasmi et.al.,(2015) "Bitcoin-based decentralized carbon emissions trading infrastructure model," *Syst. Eng.*, vol. 18, no. 2, pp. 115–130, DOI:10.1002/sys.21291.
- [15] Androulaki E et.al.,(2013) Evaluating User Privacy in Bitcoin. In: Sadeghi AR. (eds) Financial Cryptography and Data Security. Lecture Notes in Computer Science, vol 7859. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39884-1_4
- [16] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
- [17] Ezhilmathi S et.al.,(2020) "Sarcasm Sentiment Detection and Classification Model on Twitter" *Bioscience Biotechnology Research Communications Special Issue, 13(6)*, Pp:147-152.
- [18] Zengyou He et.al., (2020) "Decision tree for Sequences" *IEEE transactions on Knowledge and Data Engineering*, 10.1109/TKDE.2021.3075023.
- [19] Tie Qiu et.al., (2021) "A Directed Edge Weight Prediction Model Using Decision Tree Ensembles in Industrial Internet of Things" *IEEE Transactions On Industrial Informatics*, Vol. 17, No. 3.
- [20] Yashesh Dhebar et.al., (2020) "Interpretable Rule Discovery Through Bilevel Optimization of Split-Rules of Nonlinear Decision Trees for Classification Problems" *IEEE Transactions On Cybernetics*, DOI: 10.1109/TCYB.2020.3033003.
- [21] Ashwin Paranjape., et.al (2017) " Motifs in Temporal Networks" WSDM '17: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, Pages 601–610.
- [22] Xun Yi, et al (2019) "A New Blind ECDSA Scheme for Bitcoin Transaction Anonymity" , Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security.

Memristor based CAM cell designs and analysis of their performance.

Jugi Rithanya R P¹
 School of Electronics
 Engineering
 VIT University
 Vellore, India
 rjugi@gmail.com

Pooja K²
 School of Electronics
 Engineering
 VIT University
 Vellore, India

Sakthivel R³
 School of Electronics
 Engineering
 VIT University
 Vellore, India
 Corresponding Author
 rsakthivel@vit.ac.in

poojakeyan20@gmail.com

Abstract: A resistance with memory is what is known and defined as memristor. The incidence of the two-terminal device was found in the year 1971. Memristor is a non-volatile, expandable, high speed, high packing density, and low-power device. It will hold/retain the data, even when the power goes off. It is used in various applications like programmable logic, remote sensing, neuromorphic system, and low power applications. Random Access Memory (RAM) executes search operation by means of memory address and returns the data from the location stored. Content addressable memory (CAM) return the location of searched data by associating it with the kept data. CAM is a superior category of memory, which is used by searching applications, like caches in microprocessors, and networking. CAM cells makes parallel examination over all stowed data and finds out the matched data in one cycle. In this paper, we make a relative study on the CAM cells and the consequence of implementation of Memristor in CAM cells and analyze the performance of the design. We find that 1T-IM gives an improvement of 99% when compared to other CAM models discussed.

Keywords- CAM, Binary CAM, Memristor, Write and Read, NAND, NOR.

1. INTRODUCTION

In 1971, Leon O. Chua discovered a new two terminal device called memristor. Memristor is also presented as a fourth basic circuit element after capacitor, resistor and inductor. Memristor is well-defined as a resistor which can retain memory. Figure 1 signifies the association between the basic components and memristor. It is characterized by the association between charge and magnetic flux [1]. Subsequently thirty-six years later, HP Labs created a physical memristor with assistance from the metal titanium di-oxide. The device designed is of nanoscale size. Memristor has two regions: doped (TiO_2-x) and undoped (TiO_2) region.

The resistance in undoped region is considered as R_{off} and for doped it is taken as R_{on} . When the potential is applied, the memristor holds the value. When the potential is dropped to zero, no current flows through this physical device, but it holds the value, i.e., memristor retains the data even when the power goes off. Memristor is analysed with different models namely, Non-Linear ion drift model, Linear ion drift model, VTEAM model, Simmons tunnel barrier model, and TEAM model. [2]. VTEAM model defines about the voltage controlled memristors whereas, TEAM model deliberates on current controlled memristors [3]. Memristors are used in synaptic plasticity in the biology neuronal system [4] and also in bio-inspired pattern processing [5]. Memristors can be used in stateful logic function and so it's used in analog filters. Memristors are widely used in digital memory and logical circuits [6].

Olumodeji, O. A el al., For neuromorphic system, it is used as adaptable resistance for pulse-based operation. It has its own ability to store data [7]. It can and is used in remote sensing scheme and also in microprocessor, unconventional computing. Memristor can be used as crossbar-latches and can act as a replacement for transistors [8] and in future it can be implemented in heterogeneous systems [9].

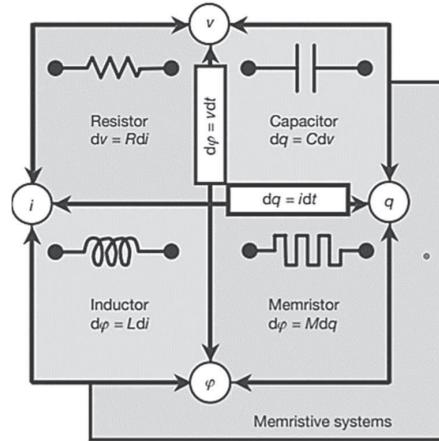


Figure 1 Relation between basic components [17]

Amamath, N el al., 10-T NOR Type CAM trails a 6-T SRAM structure with 4 additional transistors which are used to compare the stored data and the data that is searched for in the cam cell [10]. The charge division problem and short circuit current problem in NAND and NOR CAM restrict the power operation, pre-charge free CAM eliminates these drawbacks [11]. The foremost goal of pre-charge free CAM is to decrease power consumption and the speed for search application. Furthermore, higher number of searches can be achieved in lesser time. Software based search operation are slower than the hardware based CAM [12]. Conventional NAND and NOR based CAM have the disadvantages of poorer speed and high-power consumption. Bahloul el al., MCAM is considered because of its advanced speed and low power utilisation [13]. Telajala venkata mahendra el al., The hybrid architecture of three transistors and two memristors are used in 3T2M SRAM cell. These SRAM cells are used to build CAM architecture. [14]. NAND type and NOR type are very useful due to large searching performance. NOR CAM are much faster than NAND CAM, but the downfall is that it consumes higher power when compare to NAND type [15]. The advantage of the

NAND CAM is that it does not require more transistor compared to NOR CAM cell. Beside 10T NAND CAM provides complete read and write operation [16]. Yatagiri, N. H el al., The drawback in NAND CAM is that the searching delay is high and charge problem that will occur at pass transistor. The 16 cross 16 Content addressable memory arrays by means of pre-charge free reason and better-quality project to sense match data are assembled by means of the FinFET 18nm tech. In association to traditional CAM, the Precharge free CAMs provides better performance, decreases power and energy consumptions significantly[18]. Sadiq el al., the assistances of these memory cells are used in Branch Target Buffer, onchip caches, and Translation Lookaside Buffer [19]. Rouhi, S el al., this paper provides architecture of memristor-based CAMs with substantial optimization with respect to the physical layout area, power consumption, and ease of performance[20].

In this paper, we analyse the working of memristor and design various CAM cells in Cadence Virtuoso Tool. We also calculate the power consumption and timing analysis for the designs and compare the results. Section 2 discusses about the working of memristor and gives an outline on the CAM cells. Section 3 speaks on the working of CAM cells and the simulation results obtained and analyse the parameters calculated from the simulated results.

2. PROTOTYPE AND STUDY

Memristor consists of 2 layers, a layer of TiO_2 and other layer of oxygen deficient TiO_{2-x} between two platinum contacts. The two layers of TiO_2 and TiO_{2-x} can be basically represented as two variable resistance which are connected in series. Functionality of memristor device is tested by the I-V characteristics and the curve obtained is called as hysteresis curve which is obtained as shown in figure 2.1. We apply a positive voltage to the platinum contacts and initially current is zero. As the positive voltage is increased, the resistance of the device is reduced which leads to R_{ON} . When we increase the voltage more than V_{on} , the device enters in the hysteresis region. This explains that the memristor stores the data.

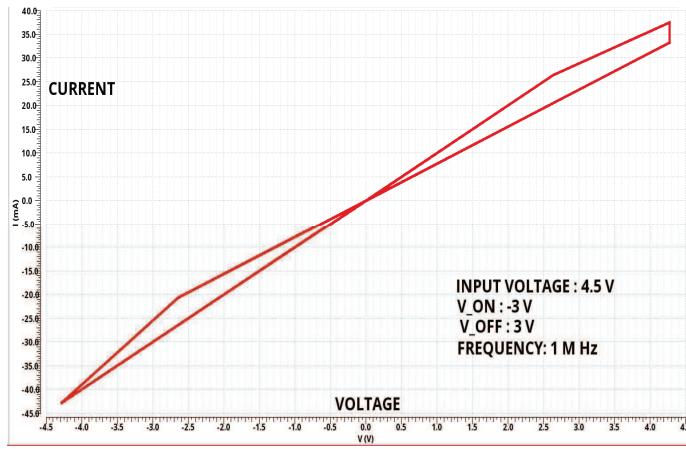


Figure 2.1 I-V Characteristics of Memristor

The device will remain in this state until a negative voltage is applied. Now, by applying negative voltage, the device resistance is increased which corresponds to the off state of the device i.e., stored data will be removed from the memory.

Content addressable memory (CAM) gives back the location of the data searched by comparing it with the stored data. CAM is a superior type of memory, which is utilised in exploratory applications, like caches on microprocessors, and computer networking. CAMs can be binary or ternary. In binary Cam, it can keep either logic-0 or 1. In ternary CAM, it can keep logic-0,1 or a don't-care. Binary CAM is used typically in cache operation, to store data and in many search operations. Ternary CAM is mostly used in application tasks. CAM makes concurrent search over all stowed words and figures out the matched data in one clock cycle. Figure 2.2 represents the block diagram of the CAM cell operation.

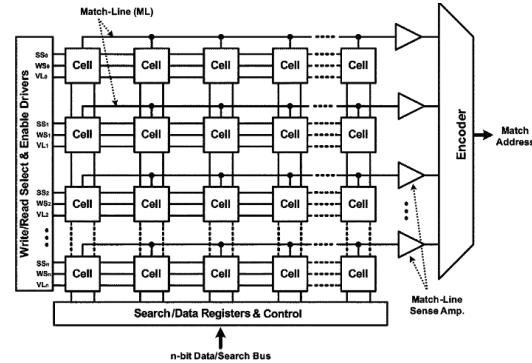


Figure 2.2 Block Diagram for CAM Cell Operation [3]

2.1 10-T NAND CAM Cell

Figure 2.3 represents the circuit diagram for 10-T NAND CAM cell. 10-T NAND consists of Word-line (WL), Bit line (BL), Bit-line bar ($BL-bar$), Q, Q-bar, Search-line, Search-line bar, Precharge Match line (ML) respectively. Word line and Word line bar are used to choose the operation. i.e. write and read operation. Bit-line and Bit-line bar is used to write the data. Search-line and Search-line bar is used to search for the stored data in the cam cell. Q and Q-bar holds the written data and it is very useful during search operation. Pre-charge match line plays a major role where the output of the match line determines whether the cell has stored the data the we searched for or not.

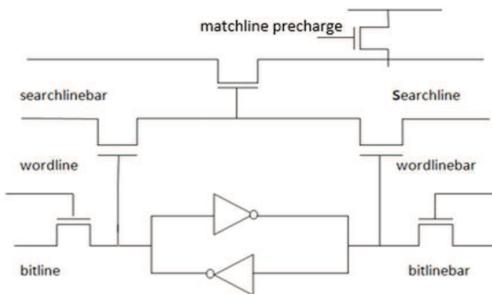


Figure 2.3 Circuit of 10-T NAND CAM Cell [3]

2.2 10-T NOR CAM Cell

Figure 2.4 represents the circuit diagram for 10-T NOR CAM cell. 10-T NOR consists of Word-line (WL), Bit line (BL), Bit-line bar (BL-bar), D,D-bar, Search-line, Search-line bar, Read-word line bar and Match line (ML) respectively. Word line and Word line bar are used to choose the operation. i.e. write and read operation. Bit-line and Bit-line bar is used to write the data. Search-line and Search-line bar is used to search for the stored data in the cam cell. D and D-bar holds the written data and it is very useful during search operation. Read-word line bar works exactly like word-line operation. Match line plays a major role where the output of the match line determines whether the cell has stored the data the we searched for or not.

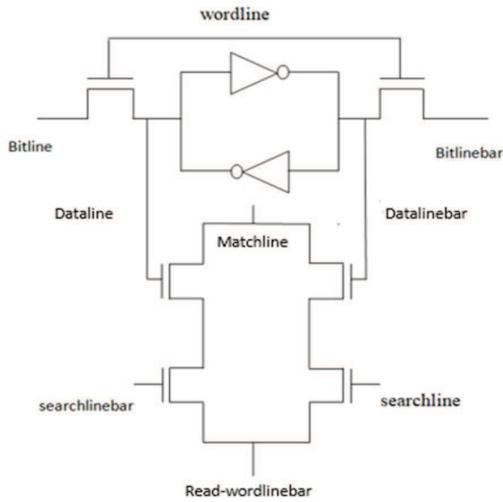


Figure 2.4 Circuit Diagram of 10T-NOR CAM Cell [3]

2.3 Pre-charge Free CAM Cell

Figure 2.5 represents the circuit diagram for pre-charge free CAM cell. In Precharge-free CAM, match line will be charged through data line when the search data matches with the stored data otherwise match line will be isolated from the data-line. Each match-line is selectively charged or not depending upon whether each bit of word is matched with the search bit or not i.e. if any bit of word matches with the stored data, then next transistor will go into saturation and charge the match-line. Pull down transistors are used to discharge this stored nodal charge before next search operation starts. When we apply search word in the Precharge free CAM, first bit of some of the CAM words will be matched and other mismatched words will be discarded from search operation. Again, in the second bit search operation some of the CAM words will be matched and other mismatched words will be discarded. Likewise bit by bit search operation continue until the end of search bit comes. As in Precharge free CAM, only few nodes are charged. So, power consumption will be largely reduced.

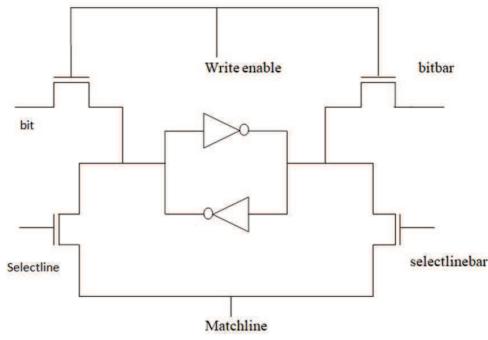


Figure 2.5 Circuit Diagram of Precharge Free CAM Cell [11]

2.4 1T-1M CAM Cell

Figure 2.6 represents the circuit diagram for 1T-1M CAM cell. It contains of Word-line and Data/search line where the gate input supply is called as Word-Line and the Data/Search-line input is given to the source of N-MOS.

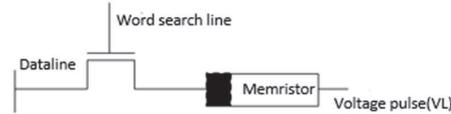


Figure 2.6 Circuit Diagram of 1T-1M CAM Cell

2.5 1T-2M CAM Cell

Figure 2.7 represents the circuit diagram for 1T-2M .The data is written into the memristor initially. The data/search line is then used to write and read the data stored in memristor.

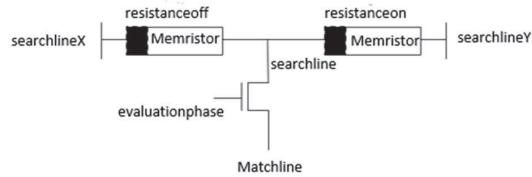


Figure 2.7 Circuit Diagram of 1T-2M CAM Cell [13]

2.6 3T-2M CAM Cell

Figure 2.8 represents the circuit diagram for 3T-2M CAM cell. In 3T2M, two memristors having reverse polarity that are linked in parallel. But during read operation they are connected in series. This altering of connections from parallel to series or vice versa are performed by the two pass transistors. Output of the OR gate is given to the gate input of the third transistor. During write operation, read signal goes to low and write signal goes to high. So, as a result of this the gate input of the third transistor goes to high state.

Then NMOS becomes in ON state and two transistors are directly connected to data-line. Now, depending upon the data input the voltage drop across memristor(*i.e* $V_D = V_{DD}/4$) positive or negative [1].

If data input is 1 then $V_D = V_{DD}$ or else $V_D = 0V$. As both the transistors are connected in opposite manner so, their resistances also change in opposite fashion. If read signal goes high in read operation, then two memristor is connected in series. Now [14], the voltage across the memristor will be

$$V_D = \left(\left(\frac{V_{DD}}{2} \right) - \left(\frac{V_{DD}}{4} \right) \right) \times \left(\frac{R_1}{R_1 + R_2} \right) + V_{DD}/4 \quad (1)$$

where the resistances R_1 and R_2 respectively are the resistances of memristor 1 and memristor 2. If R_2 is greater than R_1 then 1 is written in write operation. If R_1 is greater than R_2 that means 0 is written in write operation.

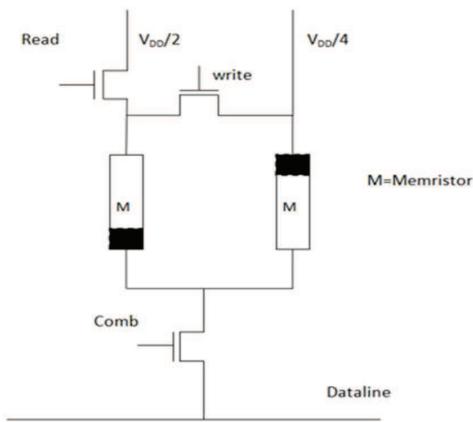


Figure 2.8 Circuit Diagram of 3T-2M CAM Cell [12]

3. SIMULATION RESULTS AND DISCUSSION

In this Section we discuss the simulation results for the above-mentioned circuits using cadence virtuoso. We also mention the working of different CAM cells and understand the performance metrics of respective designs. We have compared the results of 10T NAND Cam, 10T NOR cam, pre-charge free cam cell, 1T-1M cam cell, 1T-2M cam cell and 3T-2M cam cell. And have concluded that the 1T-1M CAM cell has the least power consumption and faster processing operation.

3.1 10-T NAND CAM Cell

During Pre-charge phase the write operation takes place where word-line is kept logic 1 and the data to be written is sent through bit-line and bit-line bar respectively. The data is stored across Q and Q bar. Match line is kept high in this phase.

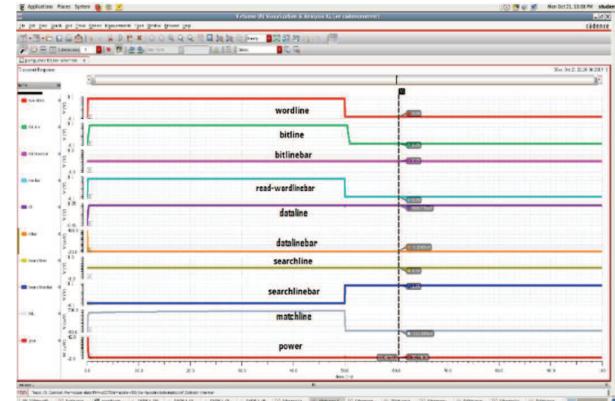


Figure 3.1 Output Waveform for 10-T NAND CAM Cell

In Evaluation phase, we keep word-line low. Data to be searched is sent through search-line and search-line bar respectively. When the data stored matches with the data searched, the pass transistor turns ON and the Match line is high. If the information stored does not match with the searched data, pass transistor does not activate and Match line is discharged to ground.

3.2 10-T NOR CAM Cell

A CAM operation has 2 phases a pre-charge phase and an evaluation phase. Throughout Pre-charge phase, the write operation takes place. First, we set the match-line to be high. Word-line and Read-Word line are kept high. The data to be written is fed into the Bit-line and Bit-line bar respectively. The data to be written is sent to the coupled inverters and the output is taken as D and D-bar. D and D bar holds the data written in the cell. Match line is high in this phase.

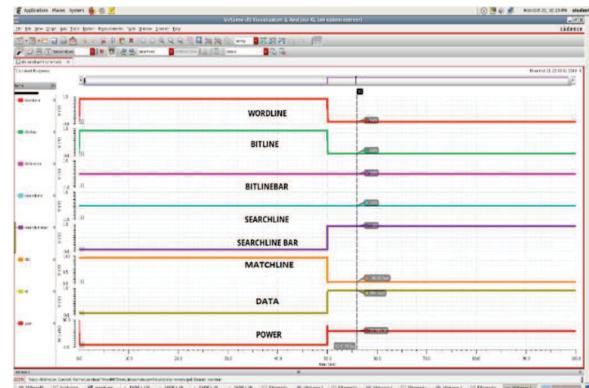


Figure 3.2 Output Waveform for 10-T NOR CAM Cell

Now, during Evaluation phase, Word-line and Read-word line bar is kept low. The data to be searched is sent into search-line and search-line bar respectively now when the data stored in D matches with the search line input, the match line will continue to stay high. This means the data stored and the searched data is same. If the data stored and the searched data does not match, match line discharges, due

to at least one pull-down transistor which connects to ground.

3.3 Pre-charge Free CAM Cell

When the search data matches with the stowed data then the match-line will be charged through the data-line. When search line goes high then it will match in bit by bit. If first bit matches with the stowed data, then match-line of the first bit goes to logic 1. The match transistor will go into saturation region. This selected match-line of first bit charges through this transistor. When the first bit matches it will search for the second bit. If second bit matches, then the selected match-line of the second bit charges. If all the bit matches with the search bit the match-line goes high. If some of the bit mismatches, then match-line will not be charged. Here we have searched for a single bit. When search-line goes high then match-line also goes high. That means searched bit present in the cell.

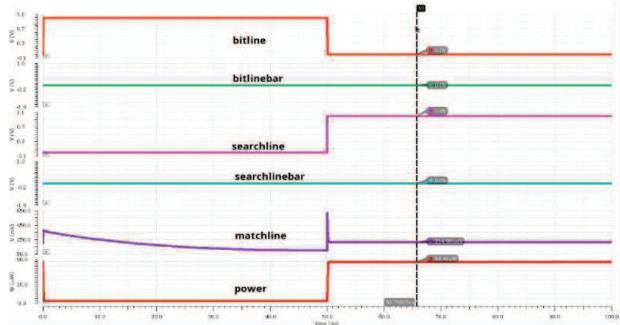


Figure 3.3 Output Waveform for Pre-charge Free CAM Cell

3.4 1T-1M CAM Cell

To write data in memristor, we maintain positive bias and negative bias when we store data '1' and data '0' respectively. Here, we keep word-line high and maintain positive bias and write data '1' respectively.

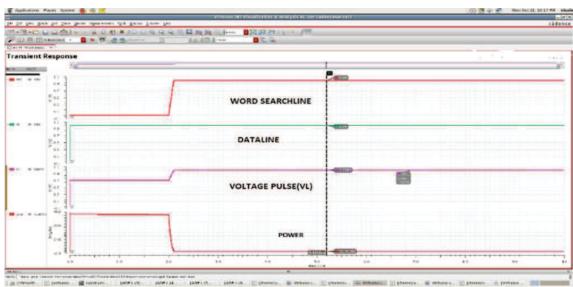


Figure 3.4 Output Waveform for 1T-1M CAM Cell

During search operation, if the search data and stowed data matches, the output line VL is high else it is low.

3.5 1T-2M CAM Cell

In the search process, two steps are required for searching a data. In pre-charge step, ML is connected to the pre-charge voltage. E is always connected to supply. In

second step, if we search for logic 0, then S_X is connected to ground and S_Y is connected to V_{DD} . If logic 1, then S_Y is connected to ground and S_X is connected to V_{DD} to search. If search data match, then voltage will keep on increase and ML wont discharged. If it mismatched, ML will be discharged to ground.

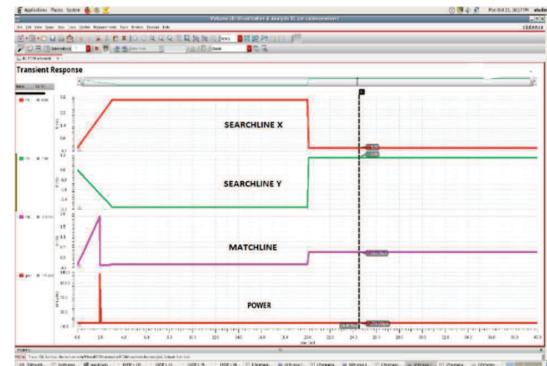


Figure 3.5 Output Waveform for 1T-2M CAM Cell

3.6 3T-2M CAM Cell

In write operation, data '1' has written in the CAM cell. Write signal goes high. Read signal goes low in write operation. So, combinational signal goes to high. As, write signal goes to high, two memristor are connected in parallel. Positive voltage drop 1V comes across the memristor. For read operation read signal goes to high and write signal goes to low. Now the two memristor is connected in series. Combinational signal also goes high. The output signal goes high. That means data found in read operation is same as the data stored in write operation.

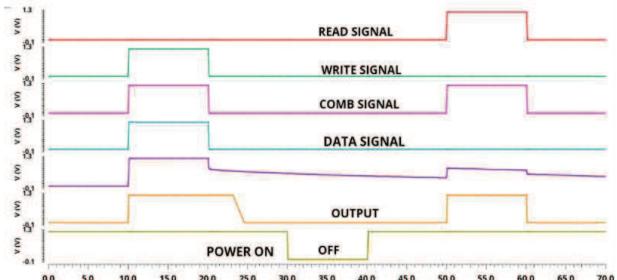


Figure 3.6 Output Waveform for 3T-2M CAM Cell

Table 1 represents the parameters that are taken for configuring memristor in CAM cell designs.

Table 1 Parameters for Memristor Configuration

DEVICE	PARAMETER	VALUE
Memristor	dt	1e-07
	ROFF	200 KΩ
	RON	100Ω
	VON	-1 V
	VOFF	+1V
	Threshold Voltage	0.5 V

From the waveforms obtained for different CAM cell design, we calculate Average time taken for CAM cell to write and search for the data and the power consumption of the CAM cells. Table 2 represents the comparison of average processing time taken for each CAM model i.e. 10-T NAND, Pre-charge Free CAM, 1T-1M, 10-T NOR, 3T-2M and 1T-1M CAM respectively. We also estimate the average power consumption for the above-mentioned CAM models.

Table 2 Comparative Analysis for Different CAM cell designs

CAM Models	Average Processing Time	Power Dissipation (nW)	Improvement in Time w.r.t 10-T NAND CAM
10-T NAND [3]	50.59 ns	16420 nW	-
Pre-charge Free [11]	50.06 ns	104.6 nW	1.24 %
10-T NOR [3]	49.96 ns	42180 nW	1.04 %
1-T 1-M	2.09e – 5ns	8.878e – 3 nW	99.9 %
1-T 2-M [13]	19.59 ns	1.93 nW	61.276 %
3-T 2-M [12]	39.12 ns	6.7 nW	22.67 %

4. CONCLUSION

We have compared the various circuits of 10T NAND Cam, 10T NOR cam, pre-charge free cam cell, 1T-1M cam cell, 1T-2M cam cell and 3T-2M cam cell using cadence virtuoso. And from the 2nd Table which explains the comparative analysis on different CAM Cell designs. From the table, we can observe that 10-T NAND consumes less power when compared to 10-T NOR, but it is bit slow when compared to 10-T NOR. Meanwhile, Pre-charge free CAM cell consumes less power when compared to 10-T NAND and 10-T NOR respectively. We can clearly say that memristor based CAM cells provides better performance when compared to other non-memristor CAM models and 1T-1M CAM cell stands out with very less power consumption and faster processing operation. When we consider memristor based CAM Models we can observe that as number of transistors increases, power consumption also increases and speed decreases. This proves that number of transistors plays a major role in power consumption and performance of the design. From the analysis, we can conclude that 1T-1M CAM model provides an improvement of 99.99% in the average time taken for one complete operation and consumes very less power when compared to other proposed models.

5. REFERENCES

- [1] Chua, L. (1971). Memristor-the missing circuit element. IEEE Transactions on circuit theory, 18(5), 507-519.
- [2] Kvatincky, S., Ramadan, M., Friedman, E. G., & Kolodny, A. (2015). VTEAM: A general model for voltage-controlled memristors. IEEE Transactions on Circuits and Systems II: Express Briefs, 62(8), 786-790
- [3] Eshraghian, K., Cho, K. R., Kavehei, O., Kang, S. K., Abbott, D., & Kang, S. M. S. (2010). Memristor MOS content addressable memory (MCAM): Hybrid architecture for future high-performance search engines. IEEE Transactions on Very Large-Scale Integration (VLSI) Systems, 19(8), 1407-1417.
- [4] Mandal, S., & Saha, A. (2016, October). Memristors act as synapses in neuromorphic architectures. In 2016 International Conference on Communication and Electronics Systems (ICCES) (pp. 1-5). IEEE.
- [5] Truong, S. N., Van Pham, K., Yang, W., & Min, K. S. (2016, October). Memristor circuits and systems for future computing and bio-inspired information processing. In 2016 IEEE Biomedical Circuits and Systems Conference (BioCAS) (pp. 456-459). IEEE.
- [6] Mokhtar, S. M. A. B., & Abdullah, W. F. H. (2014, April). Re-model fabricated memristor behavior in LT-SPICE and applied in logic circuit. In 2014 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE) (pp. 106-110). IEEE.
- [7] Olumodeji, O. A., & Gottardi, M. (2017, May). A pulse-based memristor programming circuit. In 2017 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1-4). IEEE.
- [8] Chabi, D., Wang, Z., Bennett, C., Klein, J. O., & Zhao, W. (2015). Ultrahigh density memristor neural crossbar for on-chip supervised learning. IEEE Transactions on Nanotechnology, 14(6), 954-962.
- [9] Li, H., Liu, B., Liu, X., Mao, M., Chen, Y., Wu, Q., & Qiu, Q. (2015, May). The applications of memristor devices in next-generation cortical processor designs. In 2015 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 17-20). IEEE.
- [10] Amamath, N., Vijay, H. M., Satyanarayana, S. V. V., Ramakrishnan, N. V., & Sriabibhatla, S. (2018, March). SEU sensitivity analysis of low power, pre-charge-free modified CAM cell. In 2018 4th International Conference on Devices, Circuits and Systems (ICDCS) (pp. 101-105). IEEE.
- [11] Zackriya, M., & Kittur, H. M. (2016). Precharge-free, low-power content-addressable memory. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 24(8), 2614-2621.
- [12] Sarwar, S. S., Saqueb, S. A. N., Quaiyum, F., & Rashid, A. H. U. (2013). Memristor-based nonvolatile random access memory: Hybrid architecture for low power compact memory design. IEEE Access, 1, 29-34.
- [13] Bahloul, M. A., Naous, R., & Masmoudi, M. (2017, March). Hardware emulation of memristor based ternary content addressable memory. In 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD) (pp. 446-449). IEEE.
- [14] Telalja venkata mahendra, sandeep misha," Self-Controlled High-Performance Precharge-Free Content-Addressable Memory", IEEE Transactions on Very Large-Scale Integration (VLSI) Systems Volume: 25, Issue:8, Aug. 2017
- [15] Chi-chum hung,jun-han wu,"A self-disable sense technique with differential NAND cell for CAM".2008 15TH International Conference on Devices, circuits and system.
- [16] Yang, Y., Mathew, J., Chakraborty, R. S., Ottavi, M., & Pradhan, D. K. (2016). Low cost memristor associative memory design for full and partial matching applications. IEEE Transactions on Nanotechnology, 15(3), 527-538.
- [17] Kavehei, O., Iqbal, A., Kim, Y. S., Eshraghian, K., Al-Sarawi, S. F., & Abbott, D. (2010). The fourth element: characteristics, modelling and electromagnetic theory of the memristor. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 466(2120), 2175-2202.
- [18] Yatagiri, N. H., & Sowmya, K. B. (2019, June). Low Power Self-Controlled Pre-Charge Free Content Addressable Memory. In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1225-1229). IEEE.
- [19] Sadiq, Z., & Hasan, S. (2021). MemCAM: A Hybrid Memristor-CMOS CAM Cell for On-Chip Caches. IEEE Access, 9, 21296-21305.
- [20] Rouhi, S., Mirzakuchaki, S., & Zeyrani, M. (2020, September). Optimized cell structures for memristor-based content addressable memories. In 2020 International Conference on Electrical Engineering (ICEE) (pp. 1-6). IEEE.

Loational Marginal Pricing Based Management of Congestion with Optimum Sizing of Distributed Generator using Modified ILSHADE Algorithm

Prashant

Department of Electrical Engineering,
Faculty of Engineering & Technology
Jamia Millia Islamia,
New Delhi, India
prashant.pacificcold@gmail.com

Anwar Shahzad Siddiqui

Department of Electrical Engineering,
Faculty of Engineering & Technology
Jamia Millia Islamia,
New Delhi, India
assiddiqui@jmi.ac.in

Md. Sarwar

Department of Electrical Engineering,
Faculty of Engineering & Technology
Jamia Millia Islamia,
New Delhi, India
msarwar@jmi.ac.in

Abstract—The transmission system's congestion has been aggravated by the integration of high-level renewable energy in the form of distributed generation, increasing load demand, and the ageing of the transmission network. This necessitates transmission network operators to utilize existing infrastructure facilities by employing cost-effective new transmission technologies, so as to maximise the system's potential for delivering. Congestion in the network may lead to inefficient operations or breakdowns including an outage of the linked system and network disruptions. The goal of this study is to solve the problem of congestion which may be due to sudden increase of load demand by rescheduling generators in the most efficient way possible, as well as optimal placement of distributed generator (DG) on the basis of Locational Marginal Pricing (LMP) including determination of optimal sizing of DG by the Modified IL Shade algorithm satisfying load demand. As a consequence, there will be effective congestion management, less real power loss, and an improved voltage profile, as well as an estimate of any future load that may be accommodated. On a typical IEEE-39 bus system, the validation of the proposed technique is successfully assessed in a Matlab framework through the MATPOWER tool.

Keywords—Distributed Generation, LMP, Modified IL Shade algorithm, congestion, real power loss.

I. INTRODUCTION

The rising energy consumption within power system is recognised as a serious problem for ongoing power supply system operations and expansion. As a result, power transmission congestion rises, resulting in a lower margin of network stability and security, electricity power failures, equipment damage, and a failure to properly execute energy market agreements [1]. One of the earlier ways of transmission congestion management is the nodal pricing model [2]. The independent system operator (ISO) can inject electricity at one location and receive electricity from another to alleviate congestion using this pricing approach. Authors in [3] propose a new controlling strategy for managing transmission line congestion based on nodal prices. The temperature limit of transmission lines is found in steady state using this technology, ensuring the power system's stable economic performance. The congestion effect to the power system's networks has been determined using locational marginal pricing (LMP) in [4]. Congestion management in a deregulated electrical grid is a difficult task that may be achieved by strategically placing Distributed Generators (DGs). The placement and sizing of DGs is critical for optimising the social welfare function, which includes profitability for both electricity providers and consumers[5].As, DG Resources (DER) systems are decentralised, modular, and near to the loads,

they are more adaptable. As a result, they are connected with the traditional power system in order to ease network congestion[6]. Direct techniques of line overload reduction employing generator rescheduling and load shedding have been reported in the literature [7]. To determine the capabilities of the DGs, [8] considers cost variables, load bus voltage improvements, real power flows, and power loss reduction. To achieve the positions and optimal capacities of the DGs in the electrical systems, artificial intelligence approaches such as the Fireworks Algorithm (FA) [9], Grey Wolf Optimizer (GWO) [10], Hybrid PSO (HPSO) [11], and supervised Big – Bang Crunch [12] have been used. To decrease real power loss, DGs are appropriately deployed in conjunction with network reconfiguration in [13]. Reference [6] presents a locational marginal price (LMP)-based method for allocating DG units to maximise societal welfare. Congestion management is a non-linear optimization algorithm with many variables and limitations. In the current research work, The LMP difference between the congested nodes is evaluated and utilized to evaluate the best site for DG for congestion control; higher the LMP difference, better the location for the deployment of DG. After that, to relieve congestion in the large - scale power system's transmission lines, optimum capacities of distributed generation (DG) units are injected. The proposed Modified IL-Shade algorithm having good computational efficiency is used to get the optimum DG sizes along with optimal rescheduling of the conventional generators. On standard IEEE – 39 bus test systems, the efficacy of the suggested technique is evaluated. To get the best capacities of the DG units, multi-objectives such as real power losses, investment costs, voltage variations, and line capacities are translated into a single objective and minimised. Finally, it is concluded that the suggested strategy successfully reduces congestion and is readily adaptable for tackling complicated and non-linear optimization issues in the electrical network.

The paper is represented as: section II contains discussion regarding LMP criteria for determining optimal location of DG, section III discussed the relevance of DG, section IV explains the technique of Modified IL-Shade algorithm, section V contains description of results and section VI gives the conclusion.

II.LOCATIONAL MARGINAL PRICING(LMP)

The cost of buying and selling energy at different locations across commercial energy markets, commonly referred to as Independent System Operators(ISO), is known as Locational Marginal Pricing(LMP). The

marginal cost of delivering the next increments of electrical power at a certain bus, or the added charge of sending an additional quantity of electricity through a bus, is defined as the LMP obtained within the OPF architecture[14]. Because of transmission congestion and losses, LMP may vary. Because low-cost supply cannot meet all demands when transmission gets congested, the marginal cost of energy varies by location. As a result, LMPs reflect the higher cost of delivering energy when transmission is insufficient. LMPs can differ between buses if one or maybe more transmission lines are constrained. In this situation, limited flows may preclude a cost-effective energy supply from meeting bus demands at specific sites. As a consequence, higher expensive units must be used to meet demand through other transmission lines, potentially resulting in LMP differences. The congestion cost that occurs when energy is transported from one site to another is the LMP difference between two nearby buses. Transmission losses may also have an influence on LMP variations. Marginal losses are variations in system losses as a result of incremental demand changes. The cost of marginal losses refers to the increased expenditures incurred as a result of incremental losses. These expenses are used into LMP calculations [15].

As a result, LMP is defined as follows:

$$LMP = LMP_{Energy} + LMP_{Loss} + LMP_{Congestion} \quad (1)$$

The price of power at each bus is controlled by the system limits. The locational marginal price refers to the cost disparities at each location. LMP is the marginal cost of delivering electricity to one additional MW of load at any point in the system. As a result, calculating costs depending on the location in relation to amounts of congestion power loss became an effective way for electricity trading.

III. IMPORTANCE OF DISTRIBUTED GENERATOR

Although the distributed generating(DG) units attached to local distribution networks can not only be dispatched by a central operator, they can have a major influence on power flow, voltage profile and stability including short-circuit level as well as power supply quality. The investment costs, operational cost, network reconfiguration, active and reactive power costs and power criteria are the key factors in the optimization problem [16]. Because ideally and strategically positioned DG units decreased system losses, improved system voltage profile, loadability, stability, power security, and power quality, the placement and sizing of the DG units in power network is highly important and complicated. The optimal size and position of numerous DG units were determined using a GA-based methodology that took into consideration the voltage restrictions at each node of the network to reduce system losses and electricity delivered by the grid system [17]. The optimum position of a certain number of DG units with particular total capacity was determined using a mix of PSO and GA algorithms, ensuring that the system's real power loss is reduced and its functional limitations are met [18]. When they were presented in front of a correlated optimization issue, however, there were significant defects in the efficiency

and search capability of the solution[19]. As a solution, a modified IL Shade algorithm with improved searching capabilities is being used to identify the optimal DG size. In \$-h-1, the production function of the i^{th} generator is stated as:

$$F_{Gi} = 0.5 a_{Gi} P_{Gi}^2 + b_{Gi} P_{gi} + c_{Gi} \quad (2)$$

a_{Gi} , b_{Gi} , c_{Gi} depicts fuel cost coefficients of i^{th} generator.

A quadratic cost model for DGs is utilized during this whole study.

$$C_P = (r_{DG})P^2 + (s_{DG})P + t_{DG} \quad (3)$$

r_{DG} , s_{DG} , and t_{DG} represents quadratic function coefficients given as 0.25, 40 & 0 particularly for type 2 Distributed Generator [20].

IV.METHODOLOGY OF MODIFIED IL-SHADE ALGORITHM

It's a new strategy for increasing a state-of-the-art differential evolution (DE) variant's optimization efficacy. Actually, it is the modified version of IL-Shade technique [21]. A modification of the IL-Shade method's mutation and crossover operators eliminates the issue of premature convergence and diversification of searching. For the production of the testing vector, the IL-Shade use the DE-current to pbest-1 mutation approach as shown in Eq(4):

$$\vec{vel}_{i,j} = \vec{x}_{i,j} + \vec{G}(\vec{x}_{pbest,g} - \vec{x}_{i,g}) + G(\vec{x}_{r1,g} - \vec{x}_{r2,g}) \quad (4)$$

While PSO applies [18] a modern weighted form attributable to mutation called DE-current to pbest-w-1, as shown in Eq.(5)

$$\vec{vel}_{i,j} = \vec{x}_{i,j} + G_w(\vec{x}_{pbest,g} - \vec{x}_{i,g}) + G(\vec{x}_{r1,g} - \vec{x}_{r2,g}) \quad (5)$$

The parameter G_w is altered by Eq.(6) :

$$G_w = \begin{cases} 0.7 * G, & \text{iteration current} < 0.22 * \text{iteration max} \\ 0.8 * G, & \text{iteration current} < 0.42 * \text{iteration max} \\ 1.2 * G, & \text{otherwise} \end{cases} \quad (6)$$

The modified mutation equation is now stated as Eq.(7):

$$\vec{vel}_i = \vec{x}_i + G_w(gBest - \vec{x}_i) + \vec{G}(pBest_{r1} - pBest_{r2}) \quad (7)$$

In the above equation, r_1 & r_2 demonstrates two arbitrary numbers between 1 and population size but $r_1 \neq r_2 \neq i$.

A trial vector is generated of the binomial category given by Eq.(8) as:

$$utrial_{i,j,g} = \begin{cases} vel_{i,j,g}, & \text{if Random}(0,1) \leq CR \text{ or } j = j_{Random} \\ x_{i,j,g}, & \text{otherwise} \end{cases} \quad (10)$$

for $i = 1, 2, \dots, NP$ & $j = 1, 2, \dots, D$ and $CR \in [0,1]$ indicates a crossover parameter, as well as the anticipation of creating a test vector against a mutant vector and this completes the crossover process. Eq.(11) gives the selection technique for a minimal optimum issue as:

$$\vec{x}_{i,g+1} = \begin{cases} \vec{u}_g, & \text{if } f(\vec{u}_{i,g}) \leq f(\vec{x}_{i,g}) \\ \vec{x}_{i,g}, & \text{otherwise.} \end{cases} \quad (11)$$

Where $\vec{v}_{el,i,j}$ represents the velocity vector, $vel_{i,j,g}$ represents the mutant vector; father vector is represented by $x_{i,g}$ and $j_{Random} \in \{1, 2, \dots, NP\}$ and has the responsibility of the test vector to accommodate at least one essential component of the mutant vector. The whole optimization procedure for finding optimum sizing of DG is illustrated in Fig.2.

The estimation of fitness function is done as given in Eq. (12) :

$$\text{Minimize } DG_{capacity} * DG_{cost} \quad (12)$$

Subjected to $PF_{ij}^{DG} \leq PF_{ij}^{\max}$ (13)

Wherein PF_{ij}^{DG} is power flowing after the DG has been located. & PF_{ij}^{\max} depicts maximum line capacity from one bus to the other.

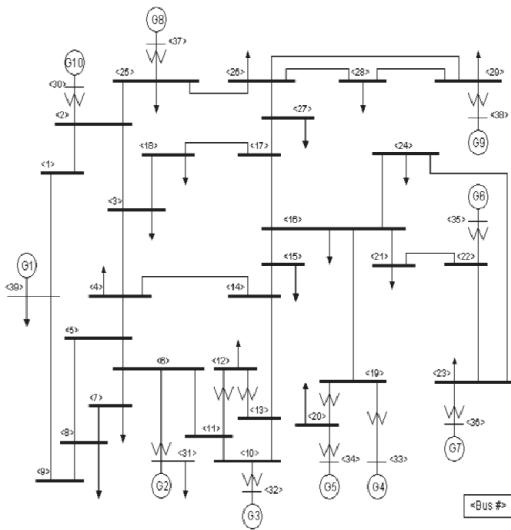


Fig.1 IEEE 39 Bus Configuration [22]

V. RESULTS AND DISCUSSION

The proposed methodology is analyzed using an IEEE radial 39 bus setup, as shown in Fig.1. 100MVA is the base. The location of 10 conventional generating units i.e. G1-G10 are positioned at following Bus no. 30, 31, 32, 33, 34, 35, 36, 37, 38, and 39. Congestion has escalated by 50% across the network, which could be due to higher load demand, building infrastructure, or other factor. Congestion is alleviated by optimizing generator output by rescheduling and injecting minimum size and number of DG at optimal places, fulfilling all limitations and satisfying load requirements.

Finding Optimum DG's Position in the Radial Network

To commence, all buses have their LMP values calculated and then LMP difference between the buses are determined. The greater the LMP difference, the more congested that node is and hence the best position for DG deployment.

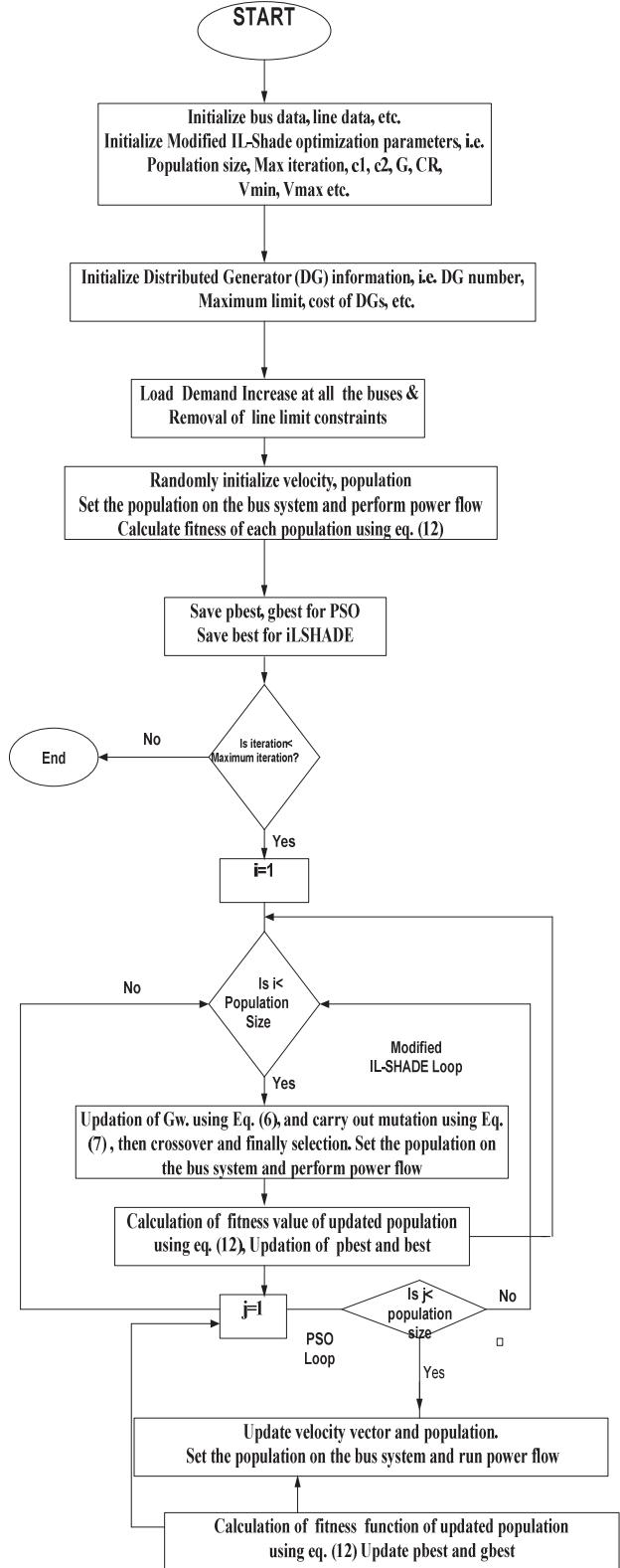


Fig.2 Optimal sizing of DG by Modified IL-SHADE algorithm

TABLE 1. LMP difference values at connecting buses

S. No	From Bus	To Bus	LMP from bus(\$-MWh-1)	LMP to bus (\$-MWh-1)	LMP Difference(\$-MWh-1)
1	1	2	13.71	114.1	100.421
2	16	19	13.54	88.72	75.182
3	29	38	13.38	85.87	72.496
4	19	33	13.42	80.63	67.212
5	21	22	13.5	80.31	66.782
6	23	24	13.73	72.21	58.483
7	20	34	13.45	63.42	49.967
8	2	3	13.83	52.68	38.854
9	25	37	13.36	47.39	34.031
10	23	36	13.45	42.81	29.356
11	6	7	13.91	41.95	28.044
12	2	25	13.44	41.80	28.359
13	25	26	13.65	37.01	23.361
14	5	8	13.94	35.56	21.618
15	6	11	13.76	26.30	12.541
16	16	21	13.65	25.84	12.185
17	15	16	13.73	23.70	9.972
18	28	29	13.51	22.55	9.041
19	26	27	13.75	20.63	6.879
20	13	14	13.81	17.37	3.559
21	1	39	14.09	16.73	2.64
22	4	14	13.81	16.17	2.363
23	8	9	14.04	15.68	1.638
24	10	11	13.76	14.63	0.87
25	26	29	13.51	14.48	0.969
26	5	6	13.83	13.38	-0.443
27	2	30	13.73	8.72	-5.0104
28	16	17	13.76	8.55	-5.2153
29	7	8	13.94	8.19	-5.7495
30	10	13	13.75	8.14	-5.6104
31	17	18	13.81	6.58	-7.2295

Table 1 shows the estimated LMP difference values at various bus locations by using methodology described earlier, and the ideal placement for DG's are at bus nos. 1, 2 and 16 for meeting the power demand while meeting all limitations according to LMP difference values. The greater the LMP difference, the higher is the congested node and hence the most optimum locations for the placement of DG's.

Finding Optimum Size of DG's

The proposed technique is used to achieve the optimal capacity of DGs, and the results are verified using the IEEE 39 setup as illustrated in Fig.2. The following are the governing parameters:

Population Size = 35; Maximum Iteration = 100; $c_1 = 2.5$; $c_2 = 2.5$; $W_{max}=0.95$ and $W_{min}= 0.55$; $G = 0.9$; $C_R = 0.75$, $W=1$; Damp = 0.98

For the OPF solution, MATPOWER [23] is used for 51 rounds with 100 iterations per round. Round 18 had several excellent results, that are reported below:

TABLE.2 OPF Solution for Round 11

S. No	From Bus	To Bus	Line Limit(MW)	Original Power Flow (MW)	Power Flow After load increment (MW)	Power Flow improvement after DG placement (MW)
1	1	2	600	173.7	70.5430	88.6615
2	1	39	1000	76.0999	292.624	58.19043
3	2	3	500	319.914	959.2402	31.29859

4	2	25	500	244.592	1442.977	213.6294
5	2	30	900	250	3.332343	244.248
6	3	4	500	37.3396	801.783	433.2333
7	3	18	500	40.7600	799.3438	401.8209
8	4	5	600	197.449	1279.324	191.3125
9	4	14	500	265.418	7374.391	45.56023
10	5	6	1200	536.936	30547.75	80.23316
11	5	8	900	339.177	45.78695	43.86216
12	6	7	900	453.815	17726.98	45.56049
13	6	11	480	322.654	16248.79	37.23635
14	6	31	1800	668.671	412.7756	77.72762
15	7	8	900	218.755	191.2319	45.51036
16	8	9	900	34.8073	25.77897	190.7692
17	9	39	900	27.9837	85.78998	674.7539
18	10	11	600	327.901	3865.587	103.8554
19	10	13	600	322.098	370.4254	425.5182
20	10	32	900	650	4.960511	371.9352
21	12	11	500	4.05615	674.7465	21.40353
22	12	13	500	4.47384	12.08129	23.67165
23	13	14	600	317.183	1241.339	461.2947
24	14	15	600	50.3143	1285.053	591.1617
25	15	16	600	269.738	3560.342	539.0725
26	16	17	600	224.017	13032.33	12.52315
27	16	19	600	451.298	3331.556	28.08582
28	16	21	600	329.601	1606.662	19.42344
29	16	24	600	42.6800	3853.324	223.457
30	17	18	600	199.038	1801.347	235.4308
31	17	27	600	24.6405	5748.488	270.8154
32	19	20	900	174.728	375.6671	575.6415
33	19	33	900	629.105	588.8812	542.2577
34	20	34	900	505.489	496.8665	503.208
35	21	22	900	604.422	657.8004	431.5674
36	22	23	600	42.7940	263.1564	40.28433
37	22	35	900	650	543.0922	492.4772
38	23	24	600	353.839	265.6079	210.8387
39	23	36	900	558.569	783.4855	572.8328
40	25	26	600	65.4138	19.27972	329.2164
41	25	37	900	538.342	97.72767	454.6584
42	26	27	600	257.295	6.636301	120.7799
43	26	28	600	140.819	0.548676	32.85785
44	26	29	600	190.188	1.345983	39.63921
45	28	29	600	347.607	201.9695	291.1858
46	29	38	1200	824.766	718.4399	786.3821

TABLE. 3 Lines that exceeded the power flow limit following a load increment and enhanced power flow after deployment of DG

S. No	From Bus	To Bus	Line Limit (MW)	Original Power Flow (MW)	Power Flow After load Increment (MW)
1	3	2	3	500	959.2402
2	4	2	25	500	1442.977
3	6	3	4	500	801.783
4	7	3	18	500	799.3438
5	8	4	5	600	1279.324
6	9	4	14	500	7374.391
7	10	5	6	1200	30547.75
8	12	6	7	900	17726.98
9	13	6	11	480	16248.79
10	18	10	11	600	3865.587
11	21	12	11	500	674.7465

12	23	13	14	600	1241.339
13	24	14	15	600	1285.053
14	25	15	16	600	3560.342
15	26	16	17	600	13032.33
16	27	16	19	600	3331.556
17	28	16	21	600	1606.662
18	29	16	24	600	3853.324
19	30	17	18	600	1801.347
20	31	17	27	600	5748.488

TABLE.4 Optimum Scheduling of Conventional Generators in MW

Ge n-1	Ge n-2	Gen -3	Gen -4	Ge n-5	Ge n-6	Gen -7	Gen -8	Gen -9	Gen-10
9.78	0	593.56	549.54	508	496.78	580	462.86	793.46	1007.93

TABLE.5 Optimal Sizing of DG by Modified IL-Shade algorithm

Bus Number	DG No.	Best Optimum Sizing(MW) Round 18	Worst Optimum Sizing (MW) Round 29
1	DG1	9.788	134.45
2	DG2	0	129.55
16	DG3	54.082	87.21
Total DG Size		63.87	351.2208

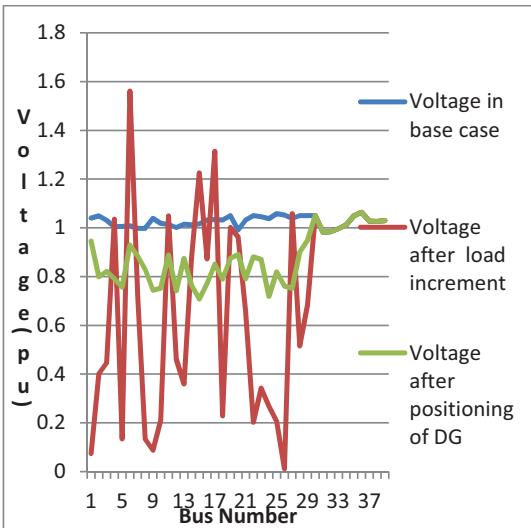


Fig.3 Voltage Profile Improvement after positioning of DG

TABLE 6.Real Power Loss in base case, following load increase and after DG positioning

Total Real Power Loss in base condition (MW)	Total Real Power Loss after increase in load (MW)	Total Real Power Loss after positioning of DG (MW)
43.64	198683.7	411.40

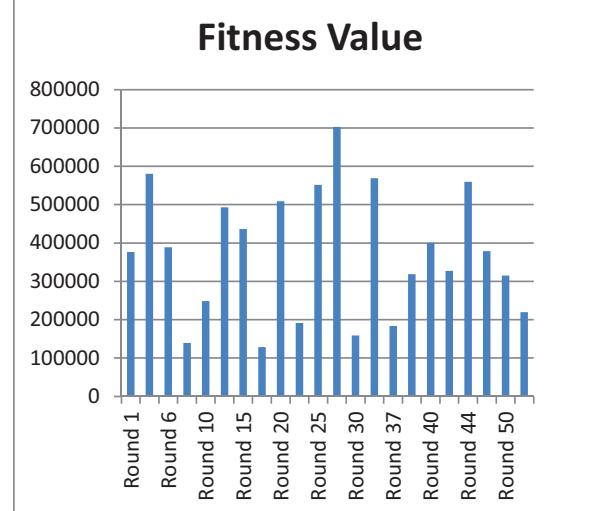


Fig.4. Sample Fitness value for different rounds of optimization.

In round 18, the best outcome of optimized power flow was reached, as can be seen in Table 2 and Table 3 depicts how lines deviate from the outflow of power restriction on increment of load and the values obtained after DG installation in the proper bus position. The optimum reschedule of conventional generators has been performed, as represented in Table.4. Fig. 3 depicts the voltage pattern progression after the DG has been installed, which was disrupted by the unexpected increase of the demand. The optimal DG sizing by Modified IL-Shade technique to be placed at optimal network locations is shown in Table.5 showing the best optimal size and worst sizing during the simulation. Table.6 shows the decrease in power loss following DG allocation, that had risen suddenly by the rapid increase in load demand and is also a substantial benefit of DG deployment. Fig.4 shows the fitness value for positioning of DG, which is best minimal in round 18, demonstrating the efficacy of the presented methodology.

VI.CONCLUSION

This study proposes the Modified IL-Shade technique as a way for resolving network congestion, which can be caused by a sudden increase in load or infrastructure development or expansion operations. Also, the growing use of distributed energy resources in power system networks has created new difficulties for network security and efficiency. The non-coordinated operation of flexible demands causes network congestion, which is a big concern. Congestion management in a deregulated energy grid is a complex challenge that can be solved by strategically placing and sizing Distributed Generators (DG). The proposed methodology is put to the test on an IEEE 39-bus network. The outcomes of the proposed approach show that rescheduling conventional generators, LMP difference-based optimum location determination, and the suggested optimization strategy for determining optimal DG sizing are successful in alleviating congestion problems, better voltage profile, reduced real power loss and prediction of any upcoming load that may be

accommodated with the existing infrastructure and facilities.

REFERENCES

- [1] E. Luo, P. Cong, H. Lu and Y. Li, "Two-Stage Hierarchical Congestion Management Method for Active Distribution Networks With Multi-Type Distributed Energy Resources," in IEEE Access, vol. 8, pp. 120309-120320, 2020, doi: 10.1109/ACCESS.2020.3005689.
- [2] D. Xiao and H. Chen, "Stochastic Up to Congestion Bidding Strategy in the Nodal Electricity Markets Considering Risk Management," in IEEE Access, vol. 8, pp. 202428-202438, 2020, doi: 10.1109/ACCESS.2020.3015025.
- [3] J. Han, A. Papavasiliou, Congestion management through topological corrections: A case study of Central Western Europe, Energy Policy, vol. 86, pp. 470-482, November, 2015.
- [4] C. Kang, Q. Chen, W. Lin, Y. Hong, Q. Xia, Z. Chen, Y. Wu, J. Xin, Zonal marginal pricing approach based on sequential network partition and congestion contribution identification, International Journal of Electrical Power & Energy Systems, vol.51, pp.321-328,2013.
- [5] S. Ghosh, S.P Ghoshal , "Optimal Sizing and Placement of Distributed Generation in a Network System," International Journal of Electrical Power & Energy Systems, Vol.32, pp.849–856, 2010.
- [6] Durga Gautam, Nadarajah Mithulanthan, "Optimal DG placement in deregulated electricity market." Electric Power Systems Research, vol. 77, issue 12, pp. 1627-1636, October, 2007.
- [7] T. K. P. Medicherla, R. Billinton and M. S. Sachdev, "Generation Rescheduling and Load Shedding to Alleviate Line Overloads-Analysis," in IEEE Transactions on Power Apparatus and Systems, vol. PAS-98, no. 6, pp. 1876-1884, Nov. 1979, doi: 10.1109/TPAS.1979.319366.
- [8] S. R. Gampa and D. Das, "Optimum placement and sizing of DGs considering average hourly variations of load," Int. J. Electr. Power Energy Syst., vol. 66, pp. 25–40, Mar. 2015.
- [9] A. Mohamed Imran, M. Kowsalya, and D. P. Kothari, "A novel integration technique for optimal network reconfiguration and distributed generation placement in power distribution networks," Int. J. Electr. Power Energy Syst., vol. 63, pp. 461–472, Dec. 2014.
- [10] U. Sultana, A. B. Khairuddin, A. S. Mokhtar, N. Zareen, and B. Sultana, "Grey wolf optimizer based placement and sizing of multiple distributed generation in the distribution system," Energy, vol. 111, pp. 525– 490, 2016.
- [11] M. M. Aman, G. B. Jasmon, A. H. A. Bakar, and H. Mokhlis, "A new approach for optimum simultaneous multi-DG distributed generation Units placement and sizing based on maximization of system loadability using HPSO (hybrid particle swarm optimization) algorithm," Energy, vol. 66, pp. 202–215, Mar. 2014.
- [12] A. Y. Abdelaziz, Y. G. Hegazy, W. El-Khattam, and M. M. Othman, "A Multi-objective Optimization for Sizing and Placement of Voltage-controlled Distributed Generation Using Supervised Big Bang-Big Crunch Method," Electr. Power Components Syst., vol. 43, no. 1, pp. 105–117, Jan. 2015.
- [13] S. Chellam, S. Kalyani, Power flow tracing based transmission congestion pricing in deregulated power markets, International Journal of Electrical Power & Energy Systems, vol.83 , 570–584,2016.
- [14] Hongli Zhang, Cong Wang, Wenhui Fan,' A new filter collaborative state transition algorithm for two-objective dynamic reactive power optimization', Tsinghua Science and Technology, Volume: 24, Issue: 1, pp:30-43, Feb. 2019.
- [15] Prashant, Siddiqui AS, Saxena A. Optimal intelligent strategic LMP Solution and Effect of DG in Deregulated System for Congestion Management. Int Trans Electr Energ Syst. 2021;e13040. <https://doi.org/10.1002/2050-7038.13040>.
- [16] Jabr, R.A., Pal, B.C.: 'Ordinal optimisation approach for locating and sizing of distributed generation', IET Gener. Transm. Distrib., 2009, 3, pp. 713–723.
- [17] Singh, R.K., Goswami, S.K.: 'Optimum allocation of distributed generations based on nodal pricing for profit, loss reduction and voltage improvement including voltage rise issue', Int. J. Electr. Power Energy Syst., 2010, 32, pp. 637–644.
- [18] Safari, A., Jahani, R., Shayanfar, H.A., Olamaei, J.: 'Optimal DG allocation in distribution network', Int. J. Electr. Electron. Eng., 2010, 4, (8), pp. 550–553.
- [19] Zwe-Lee Gaing, "A particle swarm optimization approach for optimum design of PID controller in AVR system," in IEEE Transactions on Energy Conversion, vol. 19, no. 2, pp. 384-391, June 2004.
- [20] Durga Gautam, Nadarajah Mithulanthan,"Optimal DG placement in deregulated electricity market." Electric Power Systems Research, vol. 77, issue 12, pp. 1627-1636, October, 2007.
- [21] V. Stanovov, S. Akhmedova and E. Semenkin, "LSHADE Algorithm with Rank-Based Selective Pressure Strategy for Solving CEC 2017 Benchmark Problems," 2018 IEEE Congress on Evolutionary Computation (CEC), 2018, pp. 1-8, doi: 10.1109/CEC.2018.8477977.
- [22] G. Bhatt and S. Afflajulla, "Analysis of large scale PV penetration impact on IEEE 39-Bus power system," 2017 IEEE 58th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 2017, pp. 1-6.
- [23] R. D. Zimmerman, C. E. Murillo-S anchez, and R. J. Thomas, 'Matpower: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12 {19}, Feb. 2011.

Blur Classification and Estimation of Motion Blur Parameters Using OLR

Sakthivel S

*Electronics and Communication
Amrita Vishwa Vidhyapeetham(B-tech)
Chennai, India
sakthijaya10022@gmail.com*

Akash Ram R K

*Electronics and Communication
Amrita Vishwa Vidhyapeetham(B-tech)
Chennai, India
akashram8may@gmail.com*

Sidarth Sai B

*Electronics and Communication
Amrita Vishwa Vidhyapeetham(B-tech)
Chennai, India
sidarthsai13901@gmail.com*

Ganesh Kumar C

*Electronics and Communication
Amrita Vishwa Vidhyapeetham(AP.Sr.Grd)
Chennai, India
c_ganeshkumar@ch.amrita.edu*

Abstract— The atmospheric conditions impair the quality of underwater pictures. Motion blur induced by the imaging device or the movement of the object is one of the most significant issues in underwater images. With the Artificial Intelligence and Robotics industries skyrocketing and dominating the market lately, researchers have utilized these technologies to explore and research beyond and beneath the surface of the earth and one of them being underwater image analysis and processing. But the process of underwater imaging is very complex with a lot of challenges ahead needed to be tackled. Parameters of the blurred picture must be determined to fix this in post-processing. As a result, the characteristics estimate the point spread function based on the image's spectrum. After measuring the angle and length of motion blurred pictures, the Optimized Linear Regression (OLR) approach is used to increase the parameter estimation accuracy. The data was first gathered and analyzed in real-time settings in Chennai. For underwater pictures, the suggested OLR approach outperforms the current traditional Cepstral, Hough, and Radon transform estimation methods.

Keywords— Underwater Image Processing, Motion blur, Parameter Estimation, Radon Transform, Blur classification, Linear Regression.

I. INTRODUCTION

One of the major challenges to overcome in underwater imaging is the presence of motion blur due to water currents which may lead to unsteady positioning of camera equipment and lack of controllability for accurately exposing the shot with moving subject of matter due to underwater conditions. This paper will discuss in detail about the process of countering the presence of motion blur and an effective way to classify the blur and estimate the parameters of the blurred image. Motion Blur is a frequently occurring issue, as conditions under the water make it extremely hard to capture an image that represents the subject of matter. As the water currents naturally will make the camera unsteady to a certain angle, this causes a severe motion blur in the image, making the subject poorly exposed in the captured image.

This can be resolved with sturdy equipment design to make the camera more stable but this will make the system more expensive due to additional material cost, design cost, and also this does not resolve the problem of motion blur trails in moving subjects. Even if cameras have high shutter speed, they will be highly expensive which is not cost-effective for underwater imaging. So, it is most viable to apply post-processing for a captured image using image processing techniques. There are several steps involved in the restoration

of underwater images, this paper mainly presents a method to collect appropriate data sets, data degradation analysis, identification and classification of blur type, and parameter estimation of motion blur. To make the model deployable for real-world applications, appropriate data sets are generated by applying motion blur manually to the clear underwater images with varying motion blur parameters. The extent of motion blur depends on two variables namely length, radius. Then parametric estimation is done to the generated data set. Parametric estimation for length is done with the cepstrum method and for radius is done with radon transform.

The three principal forms of blur induced by the forward scattering of light, which lowers the quality of terrestrial [1] and underwater images [2], are an atmospheric blur, motion blur, and defocus blur. Although various types of deterioration such as haze, colorcast, and noise can occur in an image, blur is a key factor in lowering the quality. As a result, underwater photos deteriorate, and latent information leads to incorrect interpretation in some applications, such as offshore excavation, dam and pipeline surveillance, species and object monitoring, and ecological research.

The approaches for detecting blur in a picture are explored and analyzed in [3]. Types of blurring will be categorized and handled based on the results of the survey. [4] presents a machine learning-based method for detecting and classifying unintended blur. The identified blur area is subjected to semantic segmentation and processing, resulting in a ground truth picture for the database. Underwater photographs, on the other hand, do not have any ground truth photos. As a result, a blind restoration procedure must be launched, followed by the use of deep learning algorithms [20] after the data has been analyzed. [5] examines the analysis of motion blurred pictures, whereas [6,7] examines estimating approaches.

II. LITERATURE SURVEY

The model stated in [15] is used to estimate the parameters of linear motion blur i.e., motion blur length and motion blur angle for a single image. The objective of this paper was to create a model that estimates the motion blur parameters with high accuracy with the combination of quick processing to get the output. From the dark lines present in the Fourier spectrum the blur direction is determined by considering the angle in between these dark lines and the vertical axis. The vertical line is extrapolated by doing a radon transform of the Coefficient all of the higher spots on the vertical lines present

in the frequency spectrum are summed up here to obtain the values of the direction or angle of motion blur.

Then for estimation of the blur length, the proposed solution for this here is NIDCT blur metric and blur metric parameters. The relationship between the blur parameters and blur metric is used to examine the blurred images with various motion lengths with constant motion direction. This data is fed into Radical Basic Function Neural Networks where each network in them is utilized for each angle. This learned network determines motion blur length when the estimated blur angle and blurriness calculated with the NIDTC blur metric is given as input.

This [16] proposes a model, that deals with the estimation of the point spread function of the motion-blurred image where the angle and the length of the blur are mainly concentrated. Also, the noise immunity of this model and the changes that are done to the noise-robust model are discussed in detail. Firstly, the angle estimation is done with the Gabor filter, where it is evident that the dominant parallel lines from the frequency spectrum are contributing to finding this parameter with ease. The convolution of the two-dimensional Gabor filter and the blurred image's spectrum is done to get the output for different orientations with having the rest of the parameters fixed. All this data extracted from different orientations are then put to a convolution with the Fourier transform of the blurred image. Then norm of every high response value is found as a result of doing convolution which in turn determines the angle of blur. To find the length parameter the summation of the Fourier coefficient's magnitude versus blur length from various images is observed and fed as an input to RBFNN. Here FFT is done to the blurred image to do a summation of Fourier coefficients and some fixed blur lengths are applied manually to some of the standard MATLAB images for training and observe a non-linear relationship followed every single image. Good noise immunity is achieved while processing the blurred images with low SNR that varies from picture to picture and the noise directly affects the vertical lines present in the frequency spectrum which is the main part of estimation. This is countered by training the RBFNN with noisy blurred images and when it comes to estimation of the angle the same previous process for estimation is repeated.

This [17] deals with a novel idea of classifying image blur based on whether they are a sharp or intentional or non-intentional blur with a classifier that detects and considers some particular hints from the blur, important pixels (mainly edges), segmentation (by labeling the objects). Other than that, this paper also discusses the removal of unintentional blurs from the pictures this is done using reducing the useless blur score and highlighting the object details present in the image, this is done using neural networks where 100 unintentional blur images with different categories are provided for training as a dataset. After which for more of a natural appeal and restore some more details from the image contour smoothed combination based on Gaussian blur is used here. The classification of blur image is done by detection of blur position in a picture, whenever the blur has occurred in the background it is safe to consider them as an intentional blur while having blur in the foreground of an image could easily be classified as an unintentional blur since

the saliency of the image is scraped away. SaNet deep learning model is used as the saliency detector, while for the blur map generation BCNet convolutional neural network is used. This model takes an input with an image segregated as 48 x 48 patches. The outputs from this model are the probabilities of sharpness, motion blur, and defocus blur. The classification here is highly influenced by the type of blur present, objects present in the image, and the places where the blur takes place. Spatial pyramid pooling is used here to determine the points where the blur distribution, saliency, and semantic segmentation are correlating and which was tested to be very close to the human's perception of intentional and unintentional blur classification. The inferred features from this pyramid are sharpness, motion blur, defocus blur, saliency, and semantics. The feature set from the pyramid is fed into the Random Forest with 1000 trees. Then the neighbor pixels are taken into account to find pixels with the similar feature as blur and the rest to isolate the blur.

The method stated in [18] is to optimize the underwater image restoration by reducing the reddish effect and image noise by doing a series of image enhancement and dehazing processes. At first, the underwater image is fed as an input where the white balancing based on grey shades and color correction is done, then the input image is fed into two different series of operations separately, where one of them is passed through a dark channel prior dehazing and contrast enhancement to reduce effects of particle scattering in the image. Then the global contrast weight, local contrast weight, saliency weight, and exposure weight, and all four weights are normalized to get one W2 norm weight and the same is done to the other input that is passed through a different series of processes. The only difference in the second series of operations from the first chain of the process is this does not implement dehazing with DCP and they also don't do contrast enhancement. The normalized weight from this chain is W1, then both of them are put through a multi-scale fusion to get the output image.

III. EXPERIMENTAL DATA COLLECTION

Figure 1 depicts the work's flow diagram. During the day, data is collected in the regions of Avadi, Mambakkam, Chenganmal, Thaivur, Thiruvananthai, ECR TNJFU, and Muttukadu in Chennai, Tamandu, India, at a maximum depth of 25 metres. The experimental data collection locations are depicted in Figure 1.

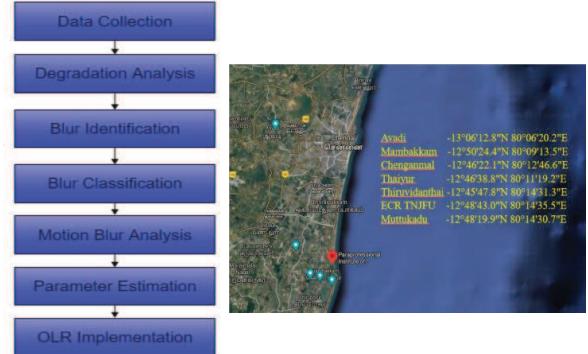


Figure 1: Workflow and places from where the Dataset Collected

The underwater drone is used to collect and study the ecological deterioration in these places. The experimental arrangement includes a drone controlled by a joystick across a 100-meter electrical tether with a topside at the other end. The data is captured in h.264 mp4 format in video mode on 720 pixels at 30 frames per second with very low latency (120 ms). The video was recorded with 6 LEDs with a total light output of 360 at a color temperature of 4000k. After that, the videos are turned into frames and processed.

The ecological circumstances in each of these seven areas vary depending on a variety of factors such as species, depth, intensity, and so on. The study of dynamic human-environment interactions using a complex systems approach is examined. The majority of the deterioration in the underwater photographs gathered is a blur. Low shutter speed of the camera, the non-uniform motion of object/camera or both, inappropriate depth of focus, mistakes in the focus of optical instruments, and atmospheric conditions can all cause blur. As a result, the blur type deterioration is restored using the appropriate post processing procedures. The example input images with distinct attributes taken at different places are shown in Figure 2.



Figure 2: Data Sets

IV. BLUR CLASSIFICATION

For underwater images, there are four types of spectrums based on the forms of blurring [10]. An irregular quadrilateral pattern for an underwater image, a disc shape for defocus blur, parallel strips with alternating light and shade for motion blur, and a four-point star-like pattern for medium blur are all shown in Figure 3. The type of blur is determined by the spectrum patterns.

A. Motion blur model

Any image acquired in a water medium will have a small angle deviation due to water flow, as seen in Figure 3, which is insignificant.

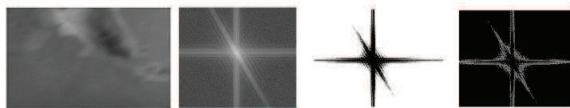


Figure 3: Motion blur

In floating media, motion blur is caused by the movement of objects, the image device, or both. Though the blur type is determined by the spectrum's shape, the point spread function must be approximated to restore the image. The two

parameters needed to estimate the motion blurred images are Length and Angle [13]. The convolution of the input picture and the degraded function with additive noise, as stated in the image formation model, leads to the identification of the restored image, which is expressed as,

$$f(x, y) = g(x, y) * h(x, y) + \eta(x, y) \quad (1)$$

$f(x, y)$ denotes the output image, $g(x, y)$ the input image, $h(x, y)$ the degradation function, and (x, y) the additive noise. Despite having the input image, the degradation function and additive noise must always be identified to restore the image. To get the degradation function for a motion blurred picture, apply the following equation [10].

$$h(x, y) = \begin{cases} \frac{1}{L} if \sqrt{x^2 + y^2} \leq \frac{L}{2} \text{ and } \frac{x}{y} = -\tan(\theta) \\ 0 \text{ otherwise} \end{cases} \quad (2)$$

where, L is the motion length and θ is the motion angle. There are parameters for each blur type. Only motion blur and defocus blur have a measurement parameter in underwater images, despite the fact that they are classified into four categories medium blur, on the other hand, resembles a hazy blur in appearance. Color restoration will thus be the best option for medium blur. Despite the fact that it has a star-shaped structure, an image normalizing approach removes the haze in underwater images and displays the irregular quadrilateral pattern in light mode. Defocus blur, restores the concentric annulus formed using radius as its parameter. In the case of a motion blurred image, the two parameters that must be assessed for restoration are length and angle. The motion blurry images are manually identified for a smaller number of images based on length and angle.

Given the multitude of frames translated from videos, manually classifying them will be a monumental task which will be nearly impossible to complete. Here the images are classified as clear images, motion-blurred images, and defocus blurred images using the VGG-16 model. The VGG-16 model [11,19] is tuned and the last layer is modified to match the desired results. With respect to the collected dataset the motion blur and defocus blur sub datasets are simulated. While for clear images, data augmentation is done some of which are shown below. The classification result is shown in Figure 4 with the precision and accuracy of 97% and 94% respectively. Motion blur restoration is the most important complex challenge to tackle in dynamic human-environment interaction, according to an analysis of underwater images.

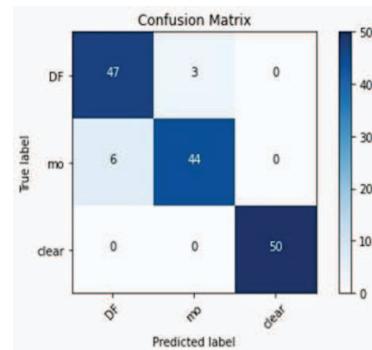


Figure 4: Classification result

V. PARAMETER ESTIMATION

The attributes of the transformation functions can be used to calculate the point spread function of blurred images. The PSF can be quantified using parametric methods such as the Hough Transform, Radon Transform, spectrum, cepstrum, steerable filters, and autocorrelation. PSF is measured and estimated here using spectrum as a baseline.

A. Angle estimation

a) *Radon transform*: The function $f(x,y)$ is represented by the Radon transform as a series of line integrals at various offsets from the origin. The sum of the intensities of the pixels in each direction of the image along the provided angles can be computed using the set of angles. The image's associated line parameters will produce a peak that is positioned for this line integral. The image matrix for the radon function is calculated along parallel paths for a given projection direction. As ray-sum is given by [12], a point of arbitrary is projected along the line $x \cos \theta + y \sin \theta = \rho$.

$$g(\rho, \theta) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) \quad (3)$$

The delta function is represented by $\delta(\cdot)$. In comparison to other transforms such as Hough and robust regression, the radon transform has the advantage of not needing the pixels of edges of the lines. The image's radon transform R must be computed first in order to determine the spectrum parallel line's direction. The motion direction is then illustrated by the position of high spots along the θ axis of R . The motion blur angle corresponds to the intersection point in the Radon transform.

B. Length estimation

a) *Cepstral Method*: The cepstrum transform can be used to separate image and blur components. Cepstrum transform of an image $f(x, y)$ is expressed as

$$c\{f(x, y)\} = F - 1(\log|F(f(x, y))|) \quad (4)$$

The sinc function from periodic patterns crosses zero in the frequency domain of uniform motion blur the negative peaks in the cepstrum domain are used to detect periodic patterns. The angle of an image is determined and then processed to estimate the blur length, as stated in the angle estimation section. The image is rotated to the expected angle for vertical projection once the angle has been calculated. Each column average is now taken in the cepstrum domain to collapse the cepstral of 2-D into 1-D [12]. Then, by locating the columns between the first negative peak and the origin, the periodicity can be determined and the blur duration can be estimated.

VI. OPTICAL LINEAR REGRESSION

After a simple haze removal using DCP, the spectrum of the motion blurred image is shown, followed by the Sobel edge detection for sharp edged spectra. The spectrum's radon transform is then shown for angle estimation where the images taken from an undersea medium, which is a non-reference source, and angle measurements are examined for a 360° cycle. As a result, the identical ratio appears in all four quadrants which are segregated with a step size of 45 degrees (as shown in Figure 5). So, for angle measurement analysis

the region in first quadrant is selected. But the measured angle differs from the origin in a non-uniform way. The estimation of the Hough transforms yields similar results, but the radon transform outperforms the Hough transform, shown in Table 1. As a result, OLR is offered as a way to improve estimation accuracy as shown in Figure 6.

After the angle estimation, the length of the blur is measured. The length of the blur is measured for up to 90 degrees and 90 pixels using the cepstrum method. As a result, the images are not square, and the length fluctuates depending on the angle. So, the length of measured angles is estimated, this approach reveals significant disparities in the lengths of comparable angles. The length estimation using the Cepstral approach is projected in Table 2.

Table 1: Angle estimation

Original Angle	Hough Transform	Radon Transform	Proposed method
5	11	9	2.83717105
10	21	18	9.76840327
15	29	26	15.92949858
20	35	33	21.32045697
25	43	40	26.71141536
30	48	46	31.33223684
35	55	51	35.18292141
40	59	56	39.03360597
45	62	61	42.88429054
50	66	65	49.0086546
55	71	69	55.52321007
60	74	72	60.40912667
65	77	75	65.29504327
70	81	78	70.18095987
75	84	81	75.06687648
80	86	84	79.95279308
85	89	87	84.83870968

Table 2: Length estimation

Angle	Actual length	Estimated Length from cepstrum	Corrected Length after OLR
10°	11	12	13.104
	23	22	25.312
	35	33	38.741
20°	18	18	20.191
	24	24	28.082
	36	35	42.548
30°	17	19	22.760
	23	24	28.232
	35	36	41.366
40°	12	11	12.011
	24	21	23.112
	36	32	35.322
50°	13	11	13.821
	25	21	24.972
	37	33	38.353
60°	14	14	15.539
	26	25	27.468
	38	36	39.397

70°	15	10	9.625
	27	20	21.578
	39	32	35.922
80°	12	11	8.469
	24	21	23.602
	36	32	40.248
90°	20	10	17.899
	26	10	17.899
	38	16	30.932

This estimation is made for all of the sample input images, where the output is the same. As a result, a single motion blurred underwater image is projected. Length can be seen varying with respect to the angle, so the angle will be calculated first. Since the length is varied with respect to the angle the lengths are also segregated into different segments, where each range is with a step size of 10 degrees are taken into account for linear regression. This provides us with low mean square error values and the accuracy of OLR-based parameter estimation of motion blurred underwater images is improved.

VII. RESULTS

The dataset for motion blur, defocus blur and clear image was manually created to increase the versatility of the training dataset. This effort helped to get the classification result that has a range of $20 \leq \theta \leq 180$ and $10 \leq L \leq 90$, this dataset was used for the training and testing of the VGG16 model created for classification of motion blur, defocus blur, and clear image. For classification of this data was achieved with an accuracy of 94%. The segregated motion-blurred images are then fed as an input for parameter estimation which are the length and angle of blur. For this, random parameters were taken in the process of testing were the length and angle are estimated using the Radon transform and the cepstrum method. Further to get more accuracy in length estimation, the estimated values are put through the Optimized Linear Regression model. This was able to achieve much lower mean square errors making this on par with many existing estimation models. The achieved Mean square error, square root of mean square error, and mean absolute error for length and angle are given in Table 3 and Table 4.

Table 3: Estimated error in angle estimation

Angle	Square root of mean square error	Mean Absolute error
45	1.394	1.217
90	0.428	0.327
135	0.911	0.785
180	1.202	1.059

Table 4: Estimated error in angle estimation

Angle	Mean Absolute error	Square root of mean square error
10	3.074	3.231
20	3.794	4.246
30	6.044	6.065
40	2.670	5.225
50	2.796	5.146
60	1.757	1.896
70	5.219	5.412
80	2.734	3.015
90	4.993	5.781

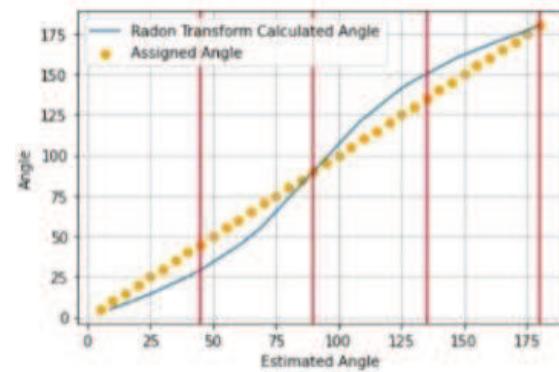


Figure 5: Radon transform results

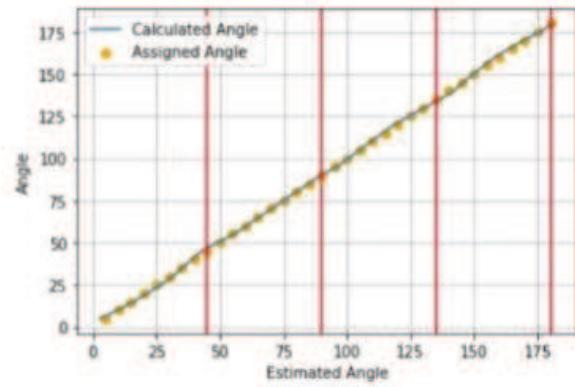


Figure 6: OLR results

VIII. CONCLUSION

The study of ecological degradation based on underwater photographs is examined. Real-time data is collected at seven different places in Chennai for this purpose. Motion blur is a complicated issue that reduces image quality. There is latent information in the motion blur. So, in order to retrieve it, the parameters for image restoration must be estimated. Estimating parameters in a dynamic undersea environment requires a more complicated approach. For which the OLR approach for parameter restoration is developed, that constantly outperforms existing transformation methods. As a result, using underwater optical pictures to analyze ecological degradations would be much simpler now.

IX. ACKNOWLEDGEMENT

We would like to thank the SSN college of engineering for providing the images over which our dataset for blur classification was based on. Also, we are obliged to acknowledge Amrita Vishwa Vidyapeetham, Chennai campus for their constant support.

X. REFERENCES

- [1] Wang, Rui, Wei Li, Rui Li, and Liang Zhang. "Automatic blur type classification via ensemble SVM." *Signal processing: image communication* 71 (2019): 24-35.
- [2] Farhadifard, Fahimeh. "Underwater image restoration: super-resolution and deblurring via sparse representation and denoising by means of marine snow removal." PhD diss., Universitat Rostock. Fakultat fur Informatik und Elektrotechnik, 2018.
- [3] Koik, Boon Tatt, and Haidi Ibrahim. "A literature survey on blur detection algorithms for digital imaging." In 2013 1st International Conference on Artificial Intelligence, Modelling and Simulation, pp. 272-277. IEEE, 2013
- [4] Huang, Rui, Mingyuan Fan, Yan Xing, and Yaobin Zou. "Image blur classification and unintentional blur removal." *IEEE Access* 7 (2019): 106327-106335.
- [5] Raj, M. Vimal, and S. Sakthivel Murugan. "Motion Deblurring Analysis for Underwater Image Restoration." In *Journal of Physics: Conference Series*, vol. 1911, no. 1, p. 012028. IOP Publishing, 2021. 12
- [6] Tiwari, Shamik, V. P. Shukla, A. K. Singh, and S. R. Biradar. "Review of motion blur estimation techniques." *Journal of Image and Graphics*, no. 4 (2013): 176-184.
- [7] Zhang, Guanyang, and Tianxiong Zheng. "Estimation Method of Blurred Parameters in Moving Blurred Image." 4 In *Journal of Physics: Conference Series*, vol. 1616, no. 1, p. 012096. IOP Publishing, 2020
- [8] Wang, Rui, Wei Li, Rui Li, and Liang Zhang. "Automatic blur type classification via ensemble SVM." *Signal 23 processing: image communication* 71 (2019): 24-35.
- [9] Qing Guan, Yunjun Wang, Bo Ping, Duanshu Li, Jiajun Du, Yu Qin, Hongtao Lu, Xiaochun Wan, Jun Xiang." Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study" *Journal of Cancer* 2019, Vol. 10 (20): 4876- 4882. Doi: 10.7150/jca.28769
- [10] Wang, Rui, Wei Li, Rui Li, and Liang Zhang. "Automatic blur type classification via ensemble SVM." *Signal 23 processing: image communication* 71 (2019): 24-35.
- [11] Qing Guan, Yunjun Wang, Bo Ping, Duanshu Li, Jiajun Du, Yu Qin, Hongtao Lu, Xiaochun Wan, Jun Xiang." Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study" *Journal of Cancer* 2019, Vol. 10 (20): 4876- 4882. Doi: 10.7150/jca.28769
- [12]. Tiwari, Shamik, V. P. Shukla, A. K. Singh, and S. R. Biradar. "Review of motion blur estimation techniques." *Journal of Image and Graphics* 1, no. 4 (2013): 176-184.
- [13] Raj, M. Vimal, and S. Sakthivel Murugan. "Motion Deblurring Analysis for Underwater Image Restoration." 45 In *Journal of Physics: Conference Series*, vol. 1911, no. 1, p. 012028. IOP Publishing, 2021
- [14]Allam Shehata Hassanein, Sherien Mohammad, Mohamed Sameer, Mohammad Ehab Ragab." A Survey on Hough Transorm, Theory, Techniques and Applications" Informatics Department, Electronics Research Institute, El-Dokki, Giza,12622, Egypt 2015.
- [15] Taiebeh Askari Javaran and Hamid Hassanpour, "Using a Blur Metric to Estimate Linear Motion Blur Parameters." Computational and Mathematical Methods in Medicine, Volume 2021.
- [16] Ratnakar Dash and Banshidhar Majhi, "Motion blur parameters estimation for image restoration." *Optik* 125 (2014) 1634–1640.
- [17] Rui Huang, Mingyuan Fan, Yan Xing, and Yaobin Zou, "Image Blur Classification and Unintentional Blur Removal." *IEEE Access* volume 7, 2019.
- [18] Yue Zhang, Fuchun Yang, and Weikai He, "An approach for underwater image enhancement based on color correction and dehazing." *International Journal of Advanced Robotic Systems* September-October 2020.
- [19] Gopakumar G, K, H. Babu, D, M., SS, G., and GR, S. Subrahmany, "Cytopathological image analysis using deep-learning networks in microfluidic microscopy", *Journal of the Optical Society of America A*, vol. 34, no. 1, pp. 111-121, 2017.
- [20] DP Nair, AM John, ERA Varshini, D Arunima, "Performance Analysis of Deep Learning Architectures for Super Resolution" *Journal of Physics: Conference Series*, 2021.

Simulation of Machine Learning Techniques to Predict Academic Performance

Praveena Chakrapani
School of Computing Sciences
Hindustan Institute of Technology and Science
Chennai, India
rp.18772016@student.hindustanuniv.ac.in

Chitradevi D
School of Computing Sciences
Hindustan Institute of Technology and Science
Chennai, India
dcdevi@hindustanuniv.ac.in

Abstract—Educational Data Mining (EDM) is that branch of Artificial Intelligence (AI) which uses a combination of Data Mining (DM) and Machine Learning (ML) techniques to make predictions on aspects specifically related to students, teachers, and, to educational institutions, in general. The goal of any educational institution is to ensure that every student gets the “right” foundational, high-quality education according to its inherent and acquired talents and potential; in order to provide the correct direction for the student’s future education and successful career. Predicting a student’s academic performance in advance, is therefore imperative to enable educational institutions and parents to take proactive decisions to steer students in the correct direction through Student Intervention Strategies. This paper simulates the most commonly used Supervised Machine Learning Algorithms that have been used for predicting Academic Performance namely, Decision Tree (DT), Random Forest (RT), K-Nearest Neighbors (KNN), Logistic Regression (LR), Naïve Bayes (NB), Gradient Boost Tree (GBT), Multi-Linear Perceptron (MLP) and Support Vector Machine (SVM). The objective of this paper is to analyze these classification algorithms for EDM in depth, in order to determine the most suitable attributes and the most efficient Machine Learning Model to accurately predict the academic performance of students and ensure academic success.

Keywords—*Machine Learning, Classification Algorithms, Supervised Learning, Educational Data Mining, Academic Performance Prediction*

I. INTRODUCTION

Modern educational institutions operate in a highly dynamic and complex environment. Analysing the wealth of educational data available and providing high-quality education has therefore become equally challenging. It is therefore of paramount importance that students’ academic performance must be predicted in an accurate and timely manner in order for educational institutions to formulate strategies and create suitable student intervention plans in order to ensure student success. Educational Data Mining (EDM) overcomes the challenges of analysing large volumes of educational data in terms of the density of data and the rich variety of attributes through Machine Learning (ML) algorithms. Scholastic, Co-scholastic, Personal, Social and Demographic characteristics are some of many segments of attributes available for Academic Performance Prediction (APP).

According to the systematic and comprehensive Literature Review by Praveena Chakrapani et al., [1], the most commonly used Supervised Classification Algorithm Methods used for Academic Prediction in the field of EDM, namely, Decision Tree (DT), Random Forest (RF), K-Nearest Neighbours (KNN), Logistic Regression (LR), Naïve Bayes (NB), Gradient Boost Tree (GBT), Multi-Linear Perceptron

(MLP) and Support Vector Machine (SVM) have been used in this research for this experimental simulation. The models have been simulated through the Machine Learning Programming Language, Python, using Python libraries (matplotlib, sklearn, numpy, pandas) and Jupyter Notebook.

Based on this simulated, comparative study, this research proposes ML Models that would ensure the most accurate prediction of a student’s academic performance, for a given data size and attribute set. The Performance metrics used for comparison are: Accuracy, Precision, Recall and F-Score. It is to be noted that in this preliminary simulation study, no boosting or fine-tuning of algorithms has been done in this initial stage of comparative performance analysis.

II. RELATED WORK

With the positive advancement in Science and Technology, there is a lot of research that is on-going in the field of EDM, particularly with respect to the Prediction of Academic Performance of students. These researches explore suitable student attributes (features) that can be used in combination with associated ML algorithms in order to devise models for accurate prediction of Academic performance.

In the study by Yahia Mohamed et al. [3], the authors have found that demographics (gender), academics (CGPA, attendance, quiz, assignments), family/personal attributes and internal assessment were the most frequently used attributes to predict students’ performance.

Aaditya Bhusal, in his study [5] uses Tensorflow techniques to determine how interaction with LMS helps the most in predicting students’ performance.

In the study by Durgesh Ugale et al. [9], the authors have found that demographics, academic, personal, family and internal assessment were the most frequently used attributes to predict students’ performance and in this study Decision Tree has provided the highest accuracy of 79%.

In the study by Vaibhav Kumar et al. [11], the authors have used school marks, continuous assessment and final examination marks in order to predict students’ performance. Here, the authors have concluded that the Gradient Boost Technique gives the highest accuracy of 98%.

In the study by Fergie Joanda Kaunang et al. [12], the authors have collected the required data using questionnaires that contained students’ demographics, previous GPA, and family background information. Decision Tree and Random Forest were applied to the students’ dataset to create the best students’ academic performance prediction model, with Decision Tree as the recommended model.

In the study by Leila Ismail et al. [13], the authors have proposed a model to predict students grades based on data availability from previous exams. The study uses various ML algorithms like Decision Tree (DT), Random Forest (RF) and Support Vector Machine (SVM) for Accuracy comparison.

Numerous studies by various authors indicate that an analysis of the various features (attributes) in their observations (number of records) has been made in order to identify the best features that could be used in the ML Algorithms in order to predict Academic Performance of students.

This research focuses on experimentally simulating the DT, RF, KNN, LR, NB, GBT, MLP and SVM models used in previous researches to solve various problems related to EDM and identifying the most suitable algorithmic model for application in Academic Performance Prediction (APP) in order to accurately predict the academic performance of students. Table I presents the most commonly used algorithms for APP, together with references to the papers that have used these methods.

TABLE I. COMMONLY USED ML TECHNIQUES FOR APP

#	ML ALGORITHM	PAPER REFERENCES
1	Decision Tree (DT)	[1, 3, 4, 6, 8, 9, 10, 11, 17, 19, 20]
2	Random Forest (RF)	[1, 6, 7, 10, 11, 12, 15]
3	K-Nearest Neighbour (KNN)	[3, 4, 9, 10, 15, 19]
4	Logistic Regression (LR)	[2, 4, 19]
5	Naïve Bayes (NB)	[5, 15, 17, 19, 20]
6	Gradient Boost Tree (GBT)	[11, 15]
7	Multi-Linear Perceptron (MLP)	[14]
8	Support Vector Machine (SVM)	[1, 2, 3, 4, 6, 7, 9, 15, 16, 20]
9	Neural Network	[17, 18, 20]

III. MATERIALS & METHODS

A. DATASET

In order to conduct the experimental simulation for evaluating the prediction accuracy of the listed ML Classifier Algorithms in Table I, for predicting academic performance, a student database from the UCI Repository has been used for this research.

TABLE II. DATA SPECIFICATIONS

Data Source	Attributes	Records (Total Students)
http://archive.ics.uci.edu/ml/machine-learning-databases/00320/	33 *	1044**

* Refer below for list of attributes considered.
** Two data sets have been merged into one dataset (395+649=1044)

List of Attributes available in the above data set:

School, Gender, Age, Address, Family Size, Parents Status, Mother's Education, Father's Education, Mother's Job, Father's Job, Reason for choosing the school, Student's Guardian, Home to School Travel Time, Weekly Study Time, Past class failures, Extra educational support, Family educational support, Extra paid classes, Extra-curricular activities, Attended nursery school, Want to take higher education, Internet access, Romantic relationship, Family relationship, Free time after school, going out with friends, daily alcohol/energy drink consumption, weekly

alcohol/energy drink consumption, health status, number of school absences, first period grade, second period grade.

Evaluation Criteria: In accordance with previous researches, the final grade is used to evaluate PASS or FAIL. For this simulation, it is considered that any grade above or equal to 10 is evaluated as PASS. Otherwise, the result is evaluated as FAIL.

B. SYSTEM FLOW ARCHITECTURE

This research runs a simulation of the following Machine Learning Techniques against the above-mentioned data set and compares the results of each algorithm with respect to the following performance metrics viz., Accuracy, Recall, Precision and F-Score: Decision Tree (DT), Random Forest (RT), K-Nearest Neighbours (KNN), Logistic Regression (LR), Naïve Bayes (NB), Gradient Boost Tree (GBT), Multi-Linear Perceptron (MLP) and Support Vector Machine (SVM).

The proposed Architectural Flow includes the following steps for each of the above 8 algorithms:

1) STEP 1

The data from the UCI Repository described in section 3.1 is collected from the data source. All the 33 attributes in the data set mentioned in Table II have been used for this experimental simulation with 1044 observations.

2) STEP 2

The collected data is validated so no incomplete observation is used for the prediction. As part of pre-processing, duplicate observations are removed and all non-numeric values are transformed using numeric values. This is done because Supervised learning Classification Machine Learning Algorithms operate on numeric values only.

3) STEP 3

The Random value 1234 has been used in order to select the training and testing data. The data set is split into Training data set and Testing data set in the ratio 70:30.

4) STEP 4

The simulation model is trained using the training data from STEP 3. All the 8 ML algorithms listed in Table I have defined methods to pass data for training. The RF Model assumed a max depth of 5, criterion as Gini and estimator value of 19. For the LR model both multinomial and OVR algorithms were considered along with maximum iteration of 1000 (random selection) to get better accuracy score and less error. The neighbour values used in KNN are: 1, 3, 5, 7, 9, 11, 13, and 25. Linear, Polynomial and Gaussian RBF kernels were used for the SVM models.

5) STEP 5

Once the model is trained, its prediction accuracy is checked using the testing data from STEP 3. Now that the model is trained, any new data can also be fed.

6) STEP 6

The results from STEP 5 across the 5 ML Algorithms are evaluated and compared. In accordance with previous researches, the evaluation criteria are based on the final grade.

The architectural flow diagram for the above process is described in Fig. 1.

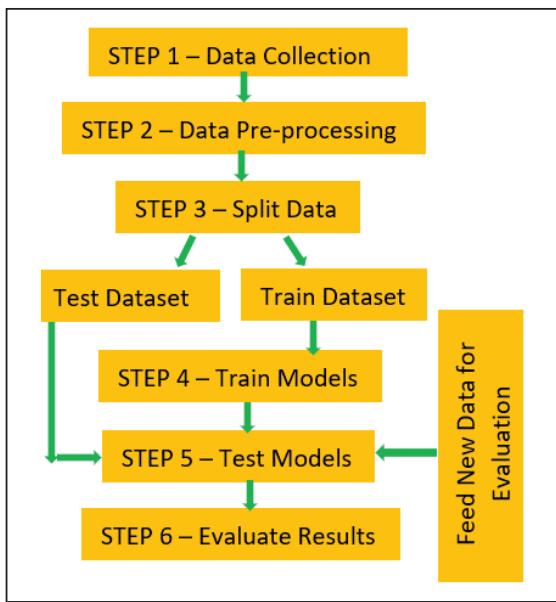


Fig. 1. Architectural Flow Diagram

IV. RESULTS

The results of this experimental simulation of the ML Models used in EDM to predict Academic Performance are listed in Figures 2 to 9 below. For this simulation, it has been considered that any grade above or equal to 10 is evaluated as PASS otherwise, the result is evaluated as FAIL.

A. DECISION TREE (DT)

Accuracy: 70%

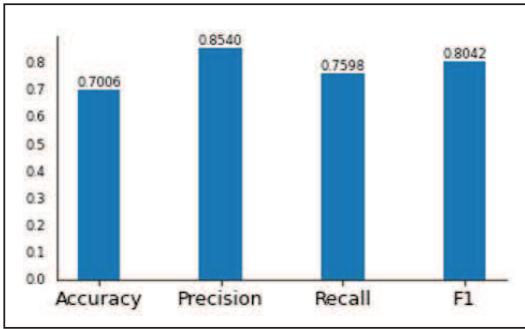


Fig. 2. Performance of the Decision Tree (DT) Algorithm

B. RANDOM FOREST (RF)

Accuracy: 81.5%

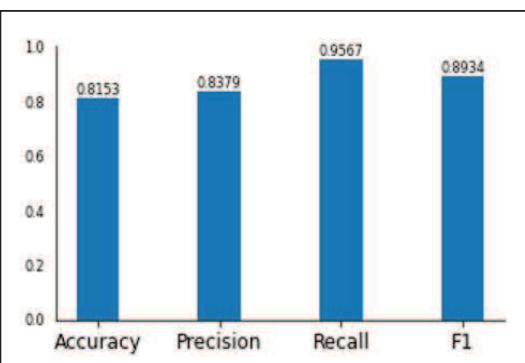


Fig. 3. Performance of the Random Forest (RF) Algorithm

C. K-NEAREST NEIGHBOUR (KNN)

Best Accuracy (25 Neighbours): 81.5%

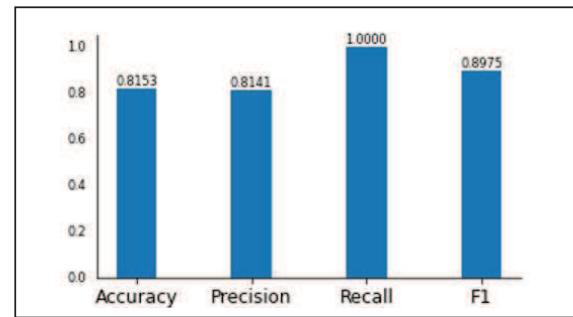


Fig. 4. Performance of K-Nearest Neighbour (KNN) Algorithm

D. LOGISTIC REGRESSION

Accuracy: 80.2%

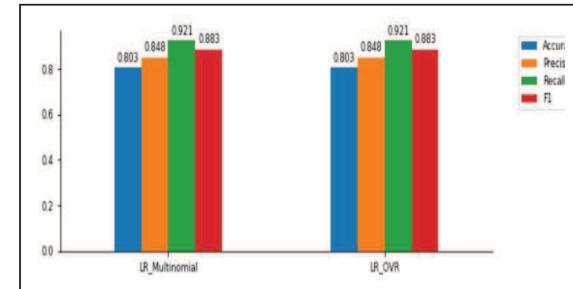


Fig. 5. Performance of the Logistic Regression (LR) Algorithm

E. NAÏVE BAYES (NB)

Accuracy: 78.9%

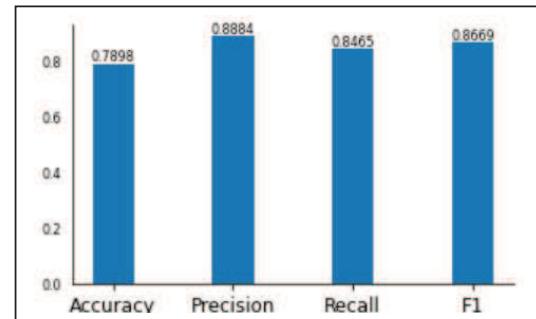


Fig. 6. Performance of the Naïve Bayes (NB) Algorithm

F. GRADIENT BOOST TREE (GBT)

Accuracy: 82.4%

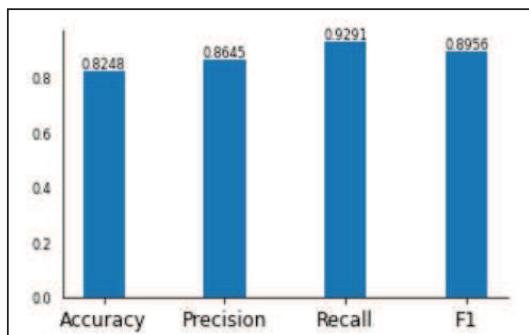


Fig. 7. Performance of the Gradient Boost Tree (GBT) Algorithm

G. MULTI-LINEAR PERCEPTRON (MLP)

Accuracy: 78.6%

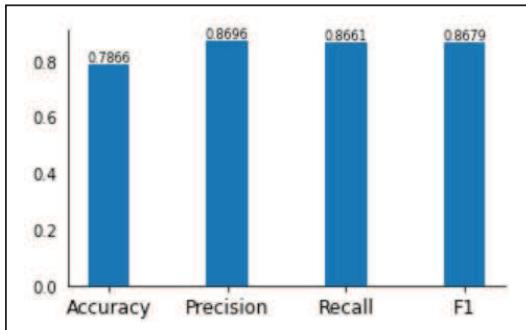


Fig. 8. Performance of the Multi-Linear Perceptron (MLP) Algorithm

H. SUPPORT VECTOR MACHINE (SVM)

Accuracy: 83.1%

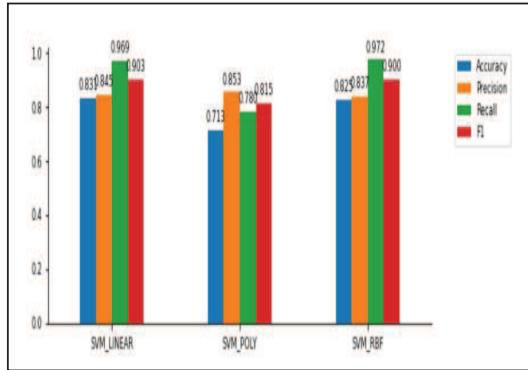


Fig. 9. Performance of the Support Vector Machine (SVM) Algorithm

V. DISCUSSION

This paper demonstrates the ability to predict student performance by using the dataset as mentioned in Table II. In this a total of 33 features were used, with data from various categories such as, academic, demographics, family and personal attributes. Using the final score, 'Pass' or 'Fail' was added to the dataset for training and testing. Any grade greater or equal to 10 is considered as PASS and rest are considered as FAIL. No feature selection or ensemble techniques were included in the simulation. In this simulation experiment, 1044 observations were used for analysis. Fig. 10 graphically illustrates the comparison of Performance metrices across the simulated algorithms.

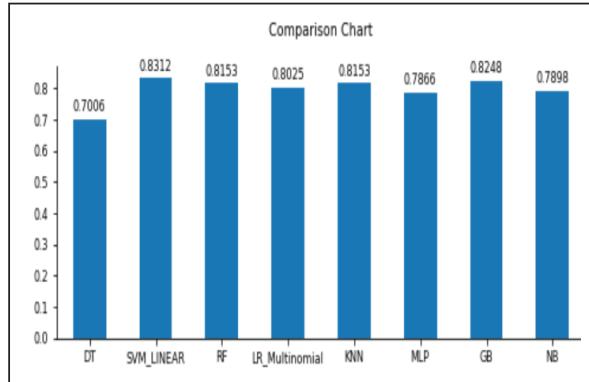


Fig. 10. Comparison of Performance Metrices Across ML Algorithms

The Performance metrices obtained during the simulation of the Academic Prediction models using DT, RF, KNN, LR, NB, GBT, MLP and SVM techniques as represented graphically in Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6, Fig. 7, Fig. 8 and Fig. 9, respectively, have been analysed in terms of:

1. Accuracy,
2. Precision,
3. Recall and,
4. F-score

in order to determine the best approach for predicting Academic Performance of students.

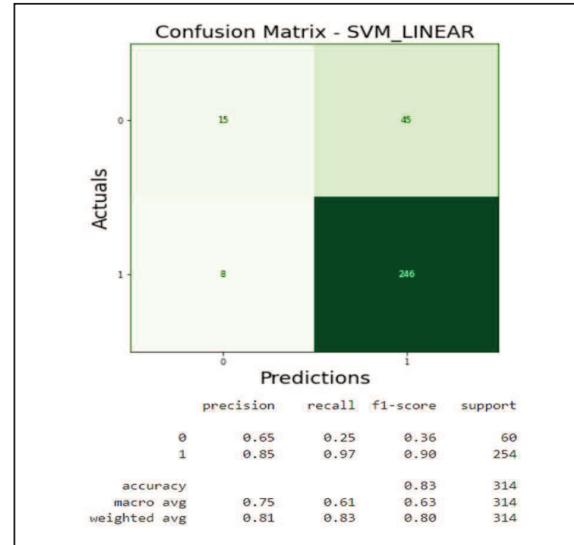


Fig. 11. Confusion Matrix for SVM Linear Algorithm

Through the analysis of this Experimental Simulation against the ML Algorithms, DT, FR, KNN, LR, NB, GBT, MLP and SVM as enumerated in Fig. 12 provides a comparison against the Performance metrices for each model, in predicting the Academic Performance of students.

	Model	Accuracy	Precision	Recall	F1
0	DT	0.700637	0.853982	0.759843	0.804167
1	SVM_LINEAR	0.831210	0.845361	0.968504	0.902752
2	SVM_POLY	0.711376	0.853448	0.779528	0.814815
3	SVM_RBF	0.824841	0.837288	0.972441	0.899818
4	RF	0.815287	0.837931	0.956693	0.893382
5	LR_Multinomial	0.802548	0.847826	0.921260	0.883019
6	LR_OVR	0.802548	0.847826	0.921260	0.883019
7	KNN_1	0.707006	0.837500	0.791339	0.813765
8	KNN_3	0.767516	0.824373	0.905512	0.863039
9	KNN_5	0.780255	0.817869	0.937008	0.873394
10	KNN_7	0.799363	0.821549	0.960630	0.885662
11	KNN_9	0.805732	0.820598	0.972441	0.890090
12	KNN_11	0.796178	0.808442	0.980315	0.886121
13	KNN_13	0.802548	0.811688	0.984252	0.889680
14	KNN	0.815287	0.814103	1.000000	0.897527
15	MLP	0.786624	0.869565	0.866142	0.867850
16	GB	0.824841	0.864469	0.929134	0.895636

Fig. 12. Outcome of the Experimental Simulation for APP

Table III summarises the best prediction in the trials for each of the ML Algorithm models using the Performance metrices Accuracy, Precision, Recall and F-Score.

TABLE III. COMPARISON OF PERFORMANCE METRICS

Model	Accuracy	Precision	Recall	F-Score
Decision Tree (DT)	0.700	0.853	0.759	0.804
Random Forest (RF)	0.815	0.837	0.972	0.899
K-Nearest Neighbour (KNN) (25)	0.815	0.814	1.0	0.897
Logistic Regression (LR)	0.80	0.847	0.921	0.883
Naïve Bayes (NB)	0.789	0.888	0.846	0.866
Gradient Boost Tree (GBT)	0.824	0.864	0.929	0.895
Multi-Linear Perceptron (MLP)	0.786	0.869	0.866	0.867
Support Vector Machine (SVM) (Linear)	0.831	0.845	0.968	0.902

Through the above simulation experiment, it has been observed that the SVM Linear Kernel Model gives the highest accuracy of 83.1% and the Decision Tree Model gives the lowest accuracy of 70% for the identical data set and identical attributes.

The Confusion Matrix metrices show that SVM Linear Kernel classifier also has the highest recall value. The ‘recall’ value is the measure of the model correctly identifying True Positives. Thus, ‘recall’ tells us how many we correctly identified as pass or fail. Mathematically,

$$\text{Recall} = \text{True Positive} / (\text{True Positive} + \text{False Negative})$$

Using the recall value, the accuracy of the relevance of prediction can be measured. For instance, it minimizes errors such as a false positive prediction resulting in no action taken whereas, in reality the student had actually failed. In the recommended SVM Linear model, the recall value of 0.97 is the highest out of all the simulated models, in addition to having the highest accuracy.

Limitations:

1. The sample dataset used is commonly available for any research and includes only the features more applicable for students in Portugal, though it can be used worldwide. Further study is required to see how these features can relate to schools and colleges in other countries where such research is in the emerging stage.

2. No feature (attribute) selection approach has been used and no data correlation has been done to find out the impact on students' academic performance. This experimental simulation indicates that selection of the right features significantly impacts the performance of the prediction metrices.

3. A small dataset of 1044 observations has been used for this experimental analysis. It has been inferred from this analysis that changes in the volume of observations leads to a different ML model for the most accurate prediction of academic performance of students.

Therefore, a greater number of experiments using various combination of attributes, data size and ML Algorithmic

Models should be conducted in order to conclude on the best model for APP (Academic Performance Prediction).

VI. CONCLUSION & FUTURE WORK

This research paper has provided an experimental simulation using various ML models that are commonly used for predicting academic performance of students. In this initial study, out of the various methods used to determine the student performance, the **SVM Linear Kernel model** proved to be best model for prediction with an accuracy score of **83.1%**.

From this experimental simulation, as tabulated in Table III and graphically represented in Fig. 10, it can be inferred that the SVM Linear Kernel model is generally a good fit for small datasets, typical of a class of students in educational institutions.

In this research, 1044 observations were used for analysis. The same approach can also be applied to predict the grades of students, in addition to PASS or FAIL criteria, with slight modifications to the data set.

The results from this experimental simulation reveals that there are specific areas in the Accuracy of APP where improvements can be made:

- (i) Identification of the correct ML technique depending on the characteristics of the data set.
- (ii) Selection of key features for improving the Accuracy of prediction of students' performance,
- (iii) Fine-tuning the model by determining the optimal values for the parameters of the model
- (iv) Application of an ensemble of boosting techniques.

REFERENCES

- [1] Praveena Chakrapani and Chitradevi D, "Academic Performance Prediction Using Machine Learning: A Comprehensive and Systematic Review", Unpublished.
- [2] Dindar Mikaeel Ahmed, Adnan Mohsin Abdulazeez, Diyar Qader Zeebaree, Falah Y. H. Ahmed, "Predicting University's Students Performance Based on Machine Learning Techniques", IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS 2021), 26 June 2021.
- [3] Yahia Mohamed, Hitham Alhussian, Gamal Alkawsi, "Predicting student's performance using machine learning methods: A systematic literature review", International Conference on Computer & Information Sciences (ICCOINS), 13-15 July 2021.
- [4] Jitendra Darji, Tulsidas Nakrani, "Enhance Student Learning Experience By Using Machine Learning To Predict Student Performance In Advance", International Research Journal of Modernization in Engineering Technology and Science, vol. 03, issue 05, May 2021.
- [5] Aaditya Bhusal, "Predicting Student's Performance Through Data Mining", unpublished.
- [6] Rimadana, M.R., Kusumawardhani, S.S., Santosa, P.I., Erwinda, M.S.F., "Predicting student academic performance using Machine Learning and Time Management Skill Data", International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2019.
- [7] S K Pushpa, T N Manjunath, T V Mrunal, Amartya Singh, C Suhas, "Class Result Prediction using Machine Learning", International Conference On Smart Technologies For Smart Nation (SmartTechCon), 2017.

- [8] Elaf Abu Amrieh, Thair Hamtini, Ibrahim Aljarah, "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods", International Journal of Database Theory and Application, Vol.9, No.8 (2016), pp.119-136, Sep 2016.
- [9] Durgesh Ugale, Jeet Pawar, Sachin Yadav, Dr. Chandrashekhar Raut, "Student Performance Prediction Using Data Mining Techniques", International Research Journal of Engineering and Technology (IRJET), vol. 07, issue 05, May 2020.
- [10] Mudasir Ashrafa, Majid Zaman, Muheet Ahmed, "An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches", International Conference on Computational Intelligence and Data Science (ICCIDDS), 2019.
- [11] Vaibhav Kumar, M.L.Garg, "Comparison of Machine Learning Models in Student Result Prediction", International Conference on Advanced Computing Networking and Informatics, pp. 439-452, 2018.
- [12] Fergie Joanda Kaunang, Reymon Rotikan, "Students' Academic Performance Prediction using Data Mining", IEEE 2018 Third International Conference on Informatics and Computing (ICIC) - Palembang, Indonesia, 2018.
- [13] Leila Ismail, Huned Materwala, and Alain Hennebelle, "Comparative Analysis of Machine Learning Models for Students' Performance Prediction", Advances in Digital Science, pp. 149-160, March 2021.
- [14] Eyman Alyahyan, Dilek Dıştegör, "Predicting academic success in higher education: literature review and best practices", International Journal of Educational Technology in Higher Education, 2020.
- [15] K. Deepika and N. Sathyaranayana, "Relief-F and Budget Tree Random Forest Based Feature Selection for Student Academic Performance Prediction", International Journal of Intelligent Engineering and Systems, 2019.
- [16] K. S. Roy, K. Roopkanth, V. U. Teja, V. Bhavana, and J. Priyanka, "Student Career Prediction Using Advanced Machine Learning Techniques", International journal of engineering and technology, vol. 7, pp. 26, 2018.
- [17] Mueen, A., Zafar, B., Manzoor, U.: Modeling and predicting students' academic performance using data mining techniques, International Journal of Modern Education and Computer Science, Nov 2016.
- [18] Hussain, S., Atallah, R., Kamsin, A., Hazarika, J.: "Classification, clustering and association rule mining in educational datasets using data mining tools", Cybernetics and Algorithms in Intelligent Systems, pp. 196-211, 2019.
- [19] W. F. W. Yaacob, S. A. M. Nasir, W. F. W. Yaacob, and N. M. Sobri, "Supervised data mining approach for predicting student performance", Vol. 16, No. 3, pp. 1584~1592, December 2019.
- [20] B. K. Francis and S. S. Babu, "Predicting academic performance of students using a hybrid data mining approach", Journal of Medical System, April 2019.

Academic Performance Prediction Using Machine Learning: A Comprehensive & Systematic Review

Praveena Chakrapani

School of Computing Sciences

Hindustan Institute of Technology and Science

Chennai, India

rp.18772016@student.hindustanuniv.ac.in

ChitraDevi D

School of Computing Sciences

Hindustan Institute of Technology and Science

Chennai, India

dcdevi@hindustanuniv.ac.in

Abstract—Educational Institutions face numerous challenges today in providing quality and student-centric education to students. Despite the huge volume of data available with educational institutions, they lack a system that monitors and analyses students' performance in order to proactively take corrective actions that would channelize the efforts of the educational institution and the student, in the correct direction by adopting Student Intervention Strategies in a timely manner. With the advent of Educational Data Mining (EDM), there is a growing awareness amongst educational institutions to utilize Data Mining (DM) and Machine Learning (ML) techniques to analyze and predict the Academic Performance of students in a reliable, sophisticated and timely manner in order to ensure academic success for every student in its institution, irrespective of the student's academic caliber. The goal of this review paper is to present a comprehensive and systematic literature review of the numerous researches done in predicting students' performance through Machine Learning techniques and assess the quality of the accuracy of predictions in a clear and crisp manner. In this review, papers published during the period 2015 to 2022 through various leading publishers have been analyzed in-depth. This comprehensive and systematic analysis clearly reveals an increasing amount of research in the area of EDM for predicting Academic Performance of students, as well as, an increasing variety and combination of ML techniques and ensemble algorithms for accurate and timely Academic Performance Prediction (APP).

Keywords—*Educational Data Mining, Machine Learning Models, Supervised Learning, Classification Algorithms, Student Academic Performance Prediction, Literature Review*

I. INTRODUCTION

All Educational Institutions consider continuous improvement to the quality of education and progressive improvement in the academic performance of its students to be of utmost importance. In order to accurately predict and achieve this goal, it is important to clearly identify the current academic level of each student and proactively apply student intervention strategies in order to attain academic success. While predicting a student's performance is a daunting and challenging task, given the huge volume and variety of student data, it is not impossible. In today's dynamic society, several factors influence a student's academic performance, rendering accurate and timely predictions of academic performance manually impossible and highly challenging even through the utilization of sophisticated AI techniques such as, Machine Learning (ML) and Data Mining (DM) used in EDM.

Most Educational Institutions of today, store a variety of data about their students in their database. Analysing this voluminous data in the traditional way is a complex and time-consuming task for even an experienced teacher and, with lack of proper tools, it is not only difficult to arrive at a prediction but with many unknown errors, it could eventually lead to

incorrect predictions that would severely impair a student's future success. This is where EDM plays a crucial role.

EDM has sophisticated methods to extract, predict and provide visual insights from raw data, through ML techniques. Analysing vast amount of data using these techniques is a boon to Educational Institutions in determining students' academic performance. Apart from this, it also educates teachers, students, academic coordinators and parents by unravelling the various attributes that contribute towards a Academic Performance Prediction (APP) and how the features can be changed in positive way in order to have a progressive impact on students' future success.

In this paper, a systematic and comprehensive literature review has been conducted on various researches done in the field of EDM for predicting academic performance, in order to provide insight into the most effective ML algorithms and attributes used for predicting students' academic performance. For this purpose, relevant articles have been identified, selected and evaluated using suitable criteria, and the findings have been integrated and summarized. From this study, it has become evident that different ML techniques reveal different prediction accuracies depending on the list of student attributes (features) used. As a result, it is important to understand that although various models are available, the appropriateness of the model is a key determiner for APP.

II. RESEARCH OBJECTIVES

The main objectives of this study are:

- Identification and understanding of various researches done in the Educational sector for APP through a comprehensive and systematic review.
- Enumeration of various attributes used for APP.
- Analysis of ML techniques used in APP.
- Categorization of the combination of attributes and MI techniques to assess the accuracy of prediction.

III. RESEARCH METHODOLOGY

A comprehensive and systematic Literature Review is imperative in order find the correct answers to the research questions, thereby paving the right direction for future research. This study follows Okoli's guide to the conduction of a systematic literature review. Fig 1 outlines the approach of Okoli's guide for systematic literature review.

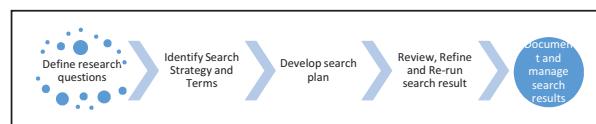


Fig. 1. Okoli's Systematic Literature Review Guide Representation

A. Define Research Questions

The first and foremost step in a comprehensive and systematic literature review is identifying the research objective and defining the research questions.

B. Search Strategy and Terms for Search

The second step defines the search strategy in choosing the best plan to accomplish the objectives. An organized and structured search strategy has been defined so that, all prior research conducted in this field is included in the process of analysis to answer the defined questions for this research.

Using Okoli's approach, extensive Google Searches along with explicit searches in authentic Computer Science journals such as IEEE, Elsevier, Springer, Science Direct, ACM and MDPI, to name a few, have been conducted.

The search terms were executed explicitly using AND / OR operators and/or truncations in search terms:

(Student Performance OR School Student Performance OR Secondary School Student Performance OR Senior Secondary School Student Performance OR Academic Performance) AND (Technique OR Method OR System OR Model OR Algorithm OR Process OR Methodology OR Procedure) AND (Educational Data Mining OR Education OR Data Mining OR Machine Learning OR Deep Learning) AND (Prediction OR Assessment OR Analysis OR Estimation OR Evaluation)

C. Search Plan

The Search Plan includes research work done during the period 2015 to 2022. Although the focus of this research is to specifically analyse ML-based Algorithms for APP, all papers relating to Machine-Learning-based Algorithms for Students' Prediction, in general, have been considered in order to conduct a comprehensive review.

D. Review, Refine and Rerun Search

Multiple iterations are required in order to find the most relevant and latest research. A Full-text Search has been conducted in order to perform a comprehensive search and to ensure that all relevant papers are included. The search has been run multiple times in order to find quality articles. Duplicate results have been eliminated.

E. Document of Results

In this step, articles which specify the Year of Publication, Source, Objective and Limitations have been tabulated in Table I and represented in Fig. 2. Results have been systematically tabulated in an Excel file for future reference.

TABLE I. NUMBER OF JOURNALS REVIEWED BY YEAR

Year	Count
2022	4
2021	11
2020	5
2019	5
2018	4
2017	2
2016	2
2015	2

Yearwise Papers Reviewed

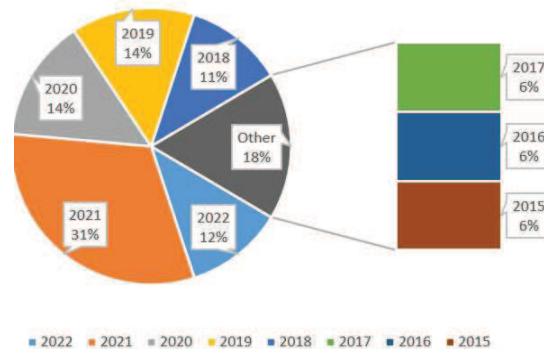


Fig. 2. Review of Research Papers by Year

IV. RESULTS

This literature review study was conducted between 2015 and 2022 and resulted in hundreds of studies during various iterative runs in order to find relevant papers. Papers refers to articles, journals, systematic literature reviews, conference papers, thesis papers and other relevant articles. This literature review selectively highlights the most relevant 35 articles out of the 100+ studies that have been reviewed in-depth. The objective is to discuss various observations that are critical to the prediction of students' academic performance.

A. Categories of Objectives Reviewed

Table II categorizes the different problems that have been researched with respect to EDM.

TABLE II. SUMMARY OF OBJECTIVES REVIEWED

Year	Papers	Count
Students' Performance in online education	[7, 10, 11, 17]	4
Students' Performance prediction based on historical data	[16, 18, 27]	3
Students' Performance in College semesters	[4, 13]	2
Students' at risk and dropout prediction	[1, 32]	2
Effect of co-curricular activities on students' performance	[14]	1
Predicting success in higher education	[6,32]	2
Predicting success in particular subjects	[21, 22, 25, 26, 35]	5
Predicting students GPA	[2, 33]	2
Academic Performance Prediction (APP)	[5, 8, 9, 12, 15, 19, 20, 23, 24, 28, 29, 30, 31, 34]	14
eLearning Textbook	[3]	1

B. Features Contributing to Students' Performance

Table III enumerates the Categories of Attributes (Features) that impact Students' Academic Performance.

TABLE III. LIST OF FEATURES

Category	Attributes
Academics	Internal Assessment, Practical Marks, Cumulative Grades, Average Grade (CGPA), Semester Subject Marks, Yearly Exam Marks, Previous Academic Marks, Assignment, Quiz, Project Results, Course Pass/Fail, Course Details
Personal	Gender, Address, Intent to pursue higher education, Attended Nursery School, Past Failures, Extra-Curricular Activities.
Family	Family Size, Parents' education, Job, Status, Income, Involvement in Education, Moral Support
Social	Internet Access, Relationship, friends, social networking, movies, outings, parties, Social Media
School	Dropout rate, Medium of education, School Reputation, School Status, Class Size, School Type, Infrastructure and facilities, etc.
Demographics	Population, Travel Time, Distance to School, Location
Online Learning	Search Activities, Face-to-Face Meetings, Discussions Viewed, Discussions Created, Assignment Viewed, Assignment Uploaded, Frequency of Comments, etc.
Co-curricular	Participate in Activities, Time Spent on Activities, etc.
eLearning	Total number of attempts to complete exercises, Number of exercises solved, Time to view the materials, Total number of hints used, Total number of exercises completed correctly, etc.

C. Machine Learning Models Recommended

Table IV and Fig. 3 summarize the recommended ML models after conducting the comprehensive and systematic review. In some of the researches, ensemble methods have been used. Ensemble methods used three main techniques namely, Boosting (BST), Bagging (BAG) and Voting (VT) to increase the accuracy score of the prediction model.

TABLE IV. RECOMMENDED MODELS

Model	References	Count
Decision Tree (DT)	[2, 13, 15, 18, 20, 24, 33]	7
Random Forest (RF)	[1, 3, 7, 10, 13, 15, 17, 23, 26, 27]	10
Support Vector Machine (SVM)	[11, 15, 30, 31]	4
Logistic Regression (LR)	[14, 32]	2
Naïve Bayes (NB)	[4, 8, 21, 28]	4
K-nearest Neighbor (KNN)	[17]	1
Artificial Neural Networks (ANN)	[12, 22]	2

Model	References	Count
Multiple Linear Perceptron (MLP)	[16, 17, 35]	3
Gradient Boost Tree (GBT)	[25]	1
Tensorflow	[9]	1
Linear Discriminant Analysis (LDA)	[16]	1
Deep Learning (DL)	[5, 19]	2
Simple Logistic (SL)	[29]	1
Rule-based Classification (RBC)	[34]	1

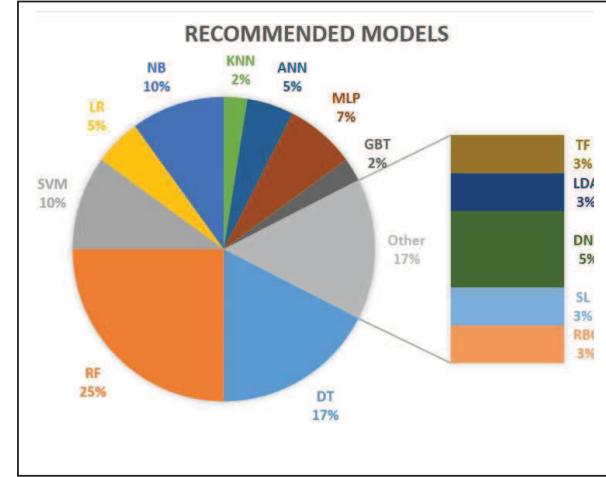


Fig. 3. Segmentation of ML Models recommended for APP

D. Model-wise Accuracy Results

Table V summarizes the best accuracy of each ML model for the chosen attributes from the most relevant research work, based on the nature of the problem definition.

TABLE V. PREDICTION ACCURACY BY MODEL

Decision Tree (DT)		
Attributes	Accuracy	Authors
Academics, Demographics, Personal	75%	[15] *for large dataset
Subject Marks, Personal	85.64%	[18]
Academics, Demographics, Personal	79%	[20]
Subject Marks	94.4%	[24]
Random Forest (RF)		
Online	85.03%	[10]
Academics, Demographics, Personal	75%	[15] *for large dataset
Academic, parents, personal, online demographics	91.5%	[23]
Academics, Demographics, Personal	93.67%	[26]
Online	85%	[7]
Academics, Parent, Online Activities	91%	[2]

Decision Tree (DT)		
Attributes	Accuracy	Authors
eLearning	91.7%	[3]
Support Vector Method (SVM)		
Academics, Demographics	70.8%	[11]
Academics, Demographics, Personal	70%	[15] *for small dataset
Academics	97%	[30]
Academics, Demographics, Personal	96%	[31]
Logistic Regression (LR)		
Co-curricular	99.52%	[14]
Semester/Subject Marks	72%	[32]
Naïve Bayes (NB)		
GPA	63.7%	[4]
Academic (Grades by Subjects)	69.67%	[8]
Demographic, Online, Personal	91.76%	[21]
Academics, Demographics, Personal	90.9%	[28]
Artificial Neural Network (ANN)		
Academic, Demographics	83%	[12]
Demographics, psychological profile, cultural, academics	77%	[22]
Multilayer Perceptron (MLP)		
Co-curricular	99.52%	[14]
Semester/Subject Marks	72%	[32]
Gradient Boost Technique (GBT)		
Marks (Academics)	98.26	[25]
Tensorflow		
Online	76%	[9]
Linear Discriminant Algorithm (LDA)		
Semester and Subject Marks (Academics)	90.74%	[16]
Deep Learning (DL)		
Academics, Personal, Family	78.2%	[19]
Personal, Online, Parent	84%	[5]
Simple Logistic (SL)		
Academics, School	95%	[29]
Rule-based Classification (RBC)		
Academics, Demographics, Personal, Family	71.3%	[34]

V. DISCUSSION

Results of the comprehensive and systematic study conducted in this research for the period 2015 to 2022, is discussed below. References to the attributes (features) and ML methods used is available in the result section.

Ghassen Ben Brahim [1], has demonstrated the ability to predict student performance during online classes by observing the interaction log files of students in the DEEDS dataset. The RF classifier with accuracy of 97% and F1-score of 97% has been recommended.

Kazeem Moses Abiodun et al. [2] has proposed data extraction techniques DT and KNN for the prediction of students' performance in exam. The authors conclude that grade is important in determining students' performance and show that DT is the effective model with a score of 91%.

Ansar Siddique et al. [6] have focused on the critical factor that affect the performance at the secondary level. A total of 16 attributes have been retained and categorized into grades such as, A+, A, B+, B, C, D and F. ML techniques MLP, PART and J48 (DT) have been used for prediction.

Sarah Alturki et al [8] have proposed a model to increase students' performance and reduce dropout rate.

Aaditya Bhusal [9] have presented a ML model to predict performance using data collected from Learning Management System. In this study, Tensorflow is used for prediction rather than other known ML models.

Tuti Purwoningsih et al. [10] have proposed a ML model to predict early academic achievement of fully online learning students. Machine learning models – RF, DT, and GBT were used and RF provided high accuracy score of 85.03%.

Derinsha Canagareddy et al. [13] have proposed an approach to predict the performance of university students and have used the Coefficient Relation, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Root Relative Squared Error (RRSE) metrices.

Shaikh Rezwan Rahman et al. [14] have proposed an approach on effects of co-curricular activities on student's academic performance. ML models - RF, VT, MLP, LR were used with LR providing 99.52% accuracy.

Fergie Joanda Kaunang et al. [26] has proposed a model to predict academic performance of Computer Science students and infer that Decision Tree is the best model with the highest accuracy value of 66.9%.

Ferda Ünal [27] has proposed the implementation of DM techniques to predict the final grades based on historical data and use an RF wrapper subset providing the highest accuracy of 93.67%.

Huda Al-Shehri et al. [31] have presented two prediction models and use correlation-based feature selection for the estimation of student's performance in the final examination and infer that SVM and KNN are most suitable.

Parneet Kaur et al. [35] have focussed on identifying slow learners among students and infer that MLP performs best with 75% accuracy.

Table VI summarises the year-wise accuracies of various MI techniques based on the problem definition of each research.

TABLE VI. YEAR-WISE ACCURACY OF MODEL BASED ON PROBLEM

Year	Problem Statement	Author	Models	Accuracy
2015	Predicting success in particular subjects	[35]	MLP, NB, DT (J48), SMO, REPTree	MLP 75%
2015	Students' Performance	[34]	DT, NB, RBC	RBC 71.3%
2016	Predicting students final GPA	[33]	DT	Not mentioned
2016	Students' at risk and dropout prediction	[32]	LR, RF, KNN	LR 72%
2017	Students' Performance Prediction	[31]	SVM, KNN	SVM 96%
2017	Students' Performance Prediction	[30]	SVM, DT, KNN, LR	SVM97%
2018	Students' Performance	[29]	Simple logistics, NB, J48	Simple logistic 95%
2018	Students' Performance	[28]	DT, LR, REPTree, RF, NB, SMO	NB 90.90%
2018	Students' performance prediction based on historical data	[27]	DT, RF, NB	RF 93.67%
2018	Predicting success in particular subjects	[26]	DT, RF, NB	RF 93.67%
2019	Predicting success in particular subjects	[25]	Linear, GBT, ANN, MLP, RF	GBT 98.26%
2019	Students' Performance	[24]	SVM, KNN, DT, RF	DT 94.4%
2019	Students' Performance	[23]	SVM, KNN, DT, RF	RF 91.5%
2019	Predicting success in particular subjects	[22]	ANN, NB, DT, LR	ANN 77%
2020	Predicting success in particular subjects	[21]	DT, NB, KNN, ET	NB 91.76%
2020	Students' Performance	[20]	DT, NB, SVM, KNN	DT 79%
2020	Students' Performance	[19]	RF, GBT, Deep Learning	DNN 78.2%
2020	Students' performance prediction based on historical data	[18]	DT, REPTree, Hoeffding	DT 85.64%
2020	Students' performance in online education	[17]	SVM, RF, NB, MLP, KNN, LR	KNN 87.5%
2020	Students' Performance Prediction	[5]	DT, RF, SVM, NB, KNN, DNN, MLP, SMO	DNN 84%
2021	Students' performance(historical data)	[16]	SVM, LDA, KNN, CART, NB, LR	LDA 90.74%
2021	Students' Performance Prediction	[15]	DT, NB, RF, SVM, ANN	SVM 70% RF 75%
2021	Effect of co-curricular activities on students' performance	[14]	VT, LR, MLP, RF	LR 99.52%
2021	Students' performance at	[13]	Nb, DT (J48), LR,	RF 100% but dataset is very less

Year	Problem Statement	Author	Models	Accuracy
	semester level or yearly basis		SVM, RF, LC	
2021	Students' Performance	[12]	NB, LR, ANN, DT	ANN 83%
2021	Students' performance in online education	[11]	SVM, DT, LR, RF, MLP	SVM 70.8%
2021	Students' performance in online education	[10]	RF, DT, GBT, AB	RF 85.03%
2021	Students' Performance	[9]	Tensorflow	Tensorflow 76%
2021	Students' Performance	[8]	DT, NB, RF	NB 69.67%
2021	Students' performance in online education	[7]	RF, NB, DT (J48), Decision Stump, OneR	RF 85%
2021	Students' performance in online education	[6]	MLP, DT, PART	MLP 98.7%
2022	Students' performance at semester level or yearly basis	[4]	DT, SVM, NB, ANN	NB 63.7%
2022	eLearning Textbook	[3]	LR, RF, KNN, SVM	RF 91.7%
2022	Predicting students final GPA	[2]	DT, KNN	DT 91%
2022	Students at risk and dropout prediction	[1]	MLP, LR, NB, SVM, RF	RF 97.4%

VI. CONCLUSION

Over several years, numerous researches have been conducted in the educational sector in order to identify student success at the outset of their academic career. Early prediction is very significant as it can aid educational institutions to develop a timely student-specific strategy and plan the implementation of necessary policies for student success. Studies have identified the best features in observations that can be used in ML models for Academic Performance Prediction (APP). Although the same ML models give different accuracy for various researches, it is clear from the present comprehensive and systematic research that features (attributes) play a critical role in determining the final accuracy, irrespective of the ML technique used. It can be concluded that same ML model gives different accuracy scores because of using varying datasets with different attributes.

Overall, this systematic and exhaustive literature review has not only provided answers for the research questions, but has also provided a strong and clear foundation for subsequent research.

VII. FUTURE WORK

While there are various researches from High school to Graduate degree, attempting to accurately predict students' performance, most are tailored to certain fixed attributes in order to arrive at a conclusion. EDM can therefore be further enhanced to predict students' performance. First, a provision of a remedial action plan to improve performance based on Academic Performance Prediction (APP). Second, the provision of a method to understand how each attribute contributes to students' success and when combined, how they

impact students' performance. Third, the provision of a dynamic model which can include / exclude features for accurate APP. Fourth, the provision of a dynamic environment, possibly through Cloud Service which can adapt, train and predict performance based on a dynamically changing set of attributes.

Educational Institutions need visual, timely and clear data to identify students who need to improve. Using ML techniques, it is also possible to provide a dynamic model that can intelligently select the best and most suitable model for prediction based on the dataset volume and prediction goal. In this way, an Educational Institution would have the flexibility to dynamically input its data and use ML techniques to intelligently and accurately predict students' academic performance in a timely manner.

REFERENCES

- [1] Ghassen Ben Brahim, "Predicting Student Performance from Online Engagement Activities Using Novel Statistical Features", Arabian Journal of Science and Engineering, Jan 2022.
- [2] Kazeem Moses Abiodun, Emmanuel Abidemi Adeniyi, Dayo Ruben Aremu, Joseph Bamidele Awotunde, Emmanuel Ogbuji, "Predicting Students Performance in Examination Using Supervised Data Mining", in book, Informatics and Intelligent Applications, pp. 63-77, Jan 2022.
- [3] Ahmed AbdElrahman, Taysir Hassan A Soliman, Ahmed I. Taloba, Mohammed F. Farghally, "A Predictive Model for Student Performance in Classrooms Using Student Interactions With an eTextbook", arXiv:2203.03713, 2022.
- [4] Muhammad Sudais, Danish Asad, "Student's Academic Performance Prediction – A Review", unpublished.
- [5] V. Vijayalakshmi, K. Venkatachalampathy, "Comparison of Predicting Student's Performance using Machine Learning Algorithms", I.J. Intelligent Systems and Applications, 2019.
- [6] Ansar Siddique, Asiya Jan, Fiaz Majeed, Adel Ibrahim Qahmash, Noorulhasan Naveed Quadri and Mohammad Osman Abdul Wahab, "Predicting Academic Performance Using an Efficient Model Based on Fusion of Classifiers Prediction Using Machine Learning Techniques", MDPI Applied Science, Dec 2021.
- [7] Kiran Fahd, Shah Jahan Miah, Khandakar Ahmed, "Predicting student performance in a blended learning environment using learning management system interaction data", in press.
- [8] Sarah Alturki, Nazik Alturki, "Using Educational Data Mining To Predict Students' Academic Performance For Applying Early Interventions", Journal of Information Technology Education: Innovations in Practice, Volume 20, pp. 121-137, 2021.
- [9] Aaditya Bhusal, "Predicting Student's Performance Through Data Mining", unpublished.
- [10] Tuti Purwoninggsih, Harry B. Santoso, Kristanti A. Puspitasari, Zainal A. Hasibuan, "Early Prediction of Students' Academic Achievement: Categorical Data from Fully Online Learning on Machine Learning Classification Algorithms", Journal of Hunan University Natural Sciences, Vol 48, No 9, 2021.
- [11] Lonia Masangu, Ashwini Jadhav, Ritesh Ajoodha, "Predicting Student Academic Performance Using Data Mining Techniques", Advances in Science, Technology and Engineering Systems Journal, vol. 6, no. 1, pp. 153-163, 2021.
- [12] Bhavesh Patel, "Performance Based Machine Learning Model to Enhance Performance of Students", Journal of University of Babylon, Pure and Applied Sciences, Vol. 27, No. 1, 2019.
- [13] Derinsha Canagareddy, Khuslendra Subaryadu, and Visham Hurbungs, "A Machine Learning Model to Predict the Performance of University Students", Smart and Sustainable Engineering for Next Generation Applications, Lecture Notes in Electrical Engineering, vol 561. Springer, Cham, 2018.
- [14] Shaikh Rezwan Rahman, Md.Asfu1 Islam, Pritidhrita Paul Akash, Masuma Parvin, Nazmun Nessa Moon, FernazNarin Nur, "Effects of co-curricular activities on student's academic performance by machine learning", International Journal of Education & Multidisciplinary Studies, ISSN 2455–2526; vol. 06, issue 03, pp. 241-254, 2017.
- [15] Leila Ismail, Huned Materwala, Alain Hennebelle, "Comparative Analysis of Machine Learning Models for Students' Performance Prediction", Advances in Digital Science, pp. 149-160, March 2021.
- [16] Jitendra Darji, Tulsidas Nakrani, "Enhance Student Learning Experience By Using Machine Learning To Predict Student Performance In Advance", International Research Journal of Modernization in Engg. Tech. & Sci., vol. 03, issue 05, May 2021.
- [17] Mohammad Noor Injadat, Abdallah Moubayed, Ali BouNassif, Abdallah Shami, "Systematic ensemble model selection approach for educational data mining", arXiv:2005.06647v1 [cs.CY], 13 May 2020.
- [18] Md. Imdadul Hoque, Abul kalam Azad, Mohammad Abu Hurayra Tuhin, Zayed Us Salehin, "University Students Result Analysis and Prediction System by Decision Tree Algorithm", Advances in Science, Technology and Engineering Systems Journal, Vol. 5, No. 3, pp. 115-122, 2020.
- [19] Samah Fakhri Aziz, "Students' Performance Evaluation Using Machine Learning Algorithms", College of Basic Education Researchers Journal Vol. 16, No. 3, Sep 2020.
- [20] Durgesh Ugale, Jeet Pawar, Sachin Yadav, Dr. Chandrashekhar Raut, "Student Performance Prediction Using Data Mining Techniques", International Research Journal of Engineering and Technology (IRJET), vol. 07, issue 05, May 2020.
- [21] Randhir Singh, Saurabh Pal, "Machine Learning Algorithms and Ensemble Technique to Improve Prediction of Students Performance", IJATCSE, Volume 9, No.3, May - June 2020.
- [22] Hussein Altabrawee, Osama Abdul Jaleel Ali, Samir Qaisar Ajm, "Predicting Students' Performance Using Machine Learning Techniques", Journal of University of Babylon, Pure and Applied Sciences, vol. 27, no.1, 2019.
- [23] Samuel-Soma M Ajibade, Nor Bahiah Binti Ahmad, Siti Mariyam Shamsuddin, "Educational Data Mining: Enhancement of Student Performance model using Ensemble Methods", Joint Conference on Green Engineering Technology & Applied Computing, 2019.
- [24] Akm Shahariar Azad Rabby, Syed Akhter Hossain, "Machine Learning Algorithm for Student's Performance Prediction", ICCCNT, Dec 2019.
- [25] Vaibhav Kumar and M. L. Garg, "Comparison of Machine Learning Models in Student Result Prediction", International Conference on Advanced Computing Networking and Informatics, pp. 439-452, 2018
- [26] Fergie Joanda Kaunang, Reymon Rotikan, "Students' Academic Performance Prediction using Data Mining", IEEE 2018 Third International Conference on Informatics and Computing (ICIC) - Palembang, Indonesia, 2018.
- [27] Ferda Ünal, "Data Mining for Student Performance Prediction in Education", in Data Mining - Methods, Applications and Systems. London, United Kingdom: IntechOpen, 2020.
- [28] R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood and K. U. Sarker, "Student Academic Performance Prediction by using Decision Tree Algorithm," 2018 4th International Conference on Computer and Information Sciences (ICCOINS), pp. 1-5, 2018.
- [29] Ismail Almuniri, Aiman Moyiad Said, "Predicting the performance of school: Case study in Sultanate of Oman", 2018 International Conference on Information and Computer Technologies (ICICT), 2018, pp. 18-21, doi: 10.1109/INFOCT.2018.8356833.
- [30] S.A. Oloruntoba, J.L. Akinode, "Student Academic Performance Prediction Using Support Vector Machine", IJESRT, Dec 2017.
- [31] Huda Al-Shehri, Amani Al-Qarni, Leena Al-Saati, Arwa Batoaq, Haifa Badukhen, Saleh Alrashed, Jamal Alhiyafi, Sunday O. Olatunji, "Student Performance Prediction Using Support Vector Machine and K-Nearest Neighbor", 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1-4, 2017.
- [32] Lovenoor Aulck, Nishant Velagapudi, Joshua Blumenstock, Jevin West, "Predicting Student Dropout in Higher Education", ArXiv abs/1606.06364, 2016.
- [33] Mashael A. Al-Barrak and Muna Al-Razgan, "Predicting Students' final GPA using decision trees", International Journal of Information and Education Technology, Vol. 6, No. 7, July 2016.
- [34] Ahmad F, Ismail N, Aziz A, "The Prediction of Students' Academic Performance Using Classification Data Mining Techniques", Applied mathematical sciences 9 : 6415-6426, 2015.
- [35] Parneet Kaur, Manpreet Singh, Gurpreet Singh Josan, "Classification and prediction based data mining algorithms to predict slow learners in education sector", Procedia Computer Science, vol. 57, pp. 500-508, 2015.

Rhythm Monitor - A Wearable for Circadian Health Monitoring

K. Sornalakshmi
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
line 5: email address or ORCID

M.Jerome Samraj
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
jm7564@srmist.edu.in

Revathi Venkataraman
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
revathin@srmist.edu.in

M.Viveka
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
vm4689@srmist.edu.in

N.Shalin
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
sn7521@srmist.edu.in

Abstract—Circadian rhythm in humans plays a vital role in people's healthy lives, but this is affected by the environment and changing lifestyles. Dysfunctions of the circadian rhythm may lead to a lot of health issues, including delayed sleep phase disorder, obesity, and many more. The cardiovascular and metabolic processes will also be affected by the circadian disruption. The present study aims to develop hardware named "Rhythm Monitor" to monitor different circadian parameters and use them to help people maintain their circadian rhythm for a healthier lifestyle. The current research is in the direction of using sensors to monitor a person's circadian function index. The results of our product help individuals by providing better recommendations to help them maintain their circadian rhythm better.

Keywords—Wearable, Sleep, light, activity, circadian function index.

I. INTRODUCTION

Circadian rhythms are 24-hour cycles [1] that coordinate the daily sleep process and affect the physical, mental, and behavioural changes of the body. These consist mainly of day and night activity and affect the daily lifestyle of all living things. Circadian rhythms play vital roles in our bodies, such as sleep induction, digestion, the release of hormones, and body temperature. Studying the biological clock and its effect on the body can help understand insomnia, jet lag, obesity, mental health disorders, and other health issues. According to surveys and studies conducted by Rush University Medical Centre, 0.13 - 3.1 % of the general population may have symptoms of DSPS (Delayed Sleep Phase Syndrome) and Irregular Circadian Rhythm contribute to DSPS [8].

The proposed work focuses on a device with built-in sensors that consistently accumulate data and information on the individual's exposure to environmental light and surrounding temperature, wrist temperature, body position, activity, and sleep [2].

The Circadian Function Index (CFI) integrates three main parameters: IS (Inter-daily stability), IV (Intra-daily variability), and RA (Relative Amplitude). CFI is the average of these three parameters. This value obtained from the formula for the Circadian function index ranges between 1 and 0. The highest value is 1 for a healthy circadian system and the lowest is 0 for an irregular circadian system [3].

Rhythm Monitor can calculate the necessary parameters to deduce the biological cycle of the respective user and estimate their daily lifestyle and hence give an outlook on how they can make changes in their routine to remain healthy.

The rhythm monitor would be perfect to assess how affected their daily routine is and how it could be improved with quicker measures taken. Rhythm Monitor helps the user find what they did wrong and gives acceptably accurate recommendations to help them identify their problems and thereby contributing to their Circadian Rhythm maintenance.

II. RELATED WORKS

Actigraph is a miniature wrist-wear device that measures limb activity, which was then used to find sleep and wake periods[4]. Actigraphy provides sustainably accurate information about an individual's sleep pattern[5]. There are many devices, including but not limited to Actiwatch and Actigraph Link available on the market which tracks the sleep pattern of an individual[6]. These devices are further advanced by providing the ability to monitor light exposure, as light exposure influences the circadian system. One of the most notable devices for monitoring circadian light exposure and activity is Actiwatch Spectrum[9]. The measuring of light exposure is imminent as the influence of light intensity on an individual cannot be generalized. It depends on the characteristics of that particular individual and how much light they are exposed to[10].

Our product, on the other hand, is a portable and lightweight device that calculates the Circadian Function Index (CFI) of each circadian parameter such as sleep, light, activity, wrist temperature, and environmental temperature. Since CFI is demonstrated as a powerful way to find the circadian rhythm of an individual, it increases the sensitivity, specificity, and accuracy of the information [7]. The Rhythm Monitor also measures light exposure and environmental temperature, which helps to identify the surrounding environment when the circadian rhythm of the individual changes. This helps the individual to have a better understanding of their own body and lifestyle.

III. PROPOSED METHOD



Fig. 1. Rhythm monitor architecture diagram

The value of CFI is calculated as the average of the three parameters IS, IV, and RA by a cloud-based circadian rhythm management system.

As evident from Figure 1, user factors such as activity, sleep, heart rate, and wrist temperature and environmental factors such as light intensity and the surrounding temperature are acquired from the “Rhythm Monitor”. The measured data are passed as inputs to the circadian rhythm management system where the necessary calculations are performed and the final CFI score can be obtained. These CFI values are displayed to the user in the app which can be accessed after the user’s authentication and thus the person can benefit from the device by making necessary changes in their lives to improve their CFI score for a healthier lifestyle. The following sensors are used to measure the circadian parameters.

A. Ambient Light Sensor

The Luminosity Sensor Breakout is used to measure the intensity of the light in terms of lux. To give a better perspective of the human eye, the sensor coalesces both visible and infrared light.

B. Contactless IR Wrist Temperature Sensor

The Non-Contact Human Body Infrared Temperature Measurement Module is used to measure the wrist temperature of the user.

C. AI-powered IoT Board

Self-learning AI smart sensors with an integrated accelerometer (3-axis) and gyroscope (3-axis) are used for detecting the activity of the user.

D. Environmental Temperature Sensor

To measure the environment temperature, we use the temperature sensor that can measure the temperature of minimum -40°C to maximum +85°C.

E. Heart Rate Sensor

It comprises a red and infrared LED, photoelectric detectors, and electronic circuits with the capability to block out unwanted light from the environment that is liable to cause interference. Pulse and blood oxygen saturation is measured by using the principle of human tissue causing different light to scatter or absorb differently according to heart rate and the blood vessels’ rhythm. The data observed from the heart rate sensor and the activity sensor is used together to find the sleep status of the user.

F. Physical configuration of the system

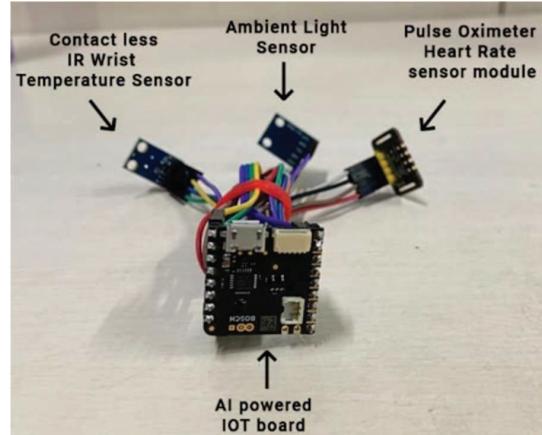


Fig. 2. Device Prototype

The Rhythm monitor system acquires the raw data from multiple sensors and normalizes it using an AI-powered IoT board and provides filtered data. These normalized values are transferred via Bluetooth to a local device for the computation of IS, IV, and RA values by the Circadian rhythm management system, which in turn provides the overall CFI value.

The Rhythm monitor system starts by acquiring the raw data from the numerous sensors available (such as contactless IR wrist temperature sensor, ambient light sensor, the pulse oximeter/ heart rate sensor, etc...). The raw data is obtained and normalized by the system. Normalizing data helps in reducing the processing time and power that is needed throughout the rest of the process. The normalized data is then sent to the local device over Bluetooth, this device will take care of the computation of the IS (Interdaily stability), IV (Intradaily variability), and RA (Relative Amplitude) values. It then plugs these values into the formula to compute the overall CFI (Circadian function index) value, which can be displayed to the user as their circadian rhythm score.

G. Circadian Rhythm Management System

The deployed model of the circadian rhythm monitor is a bundled data abstraction of the parameters acquired from the 3 important modules, which are Interdaily Stability (IS), Intradaily Variability (IV), and Relative Amplitude (RA) all together comprising the overall Circadian Function Index (CFI) value respectively.

The system acquires the necessary parameters, which are stored after further filtration and normalization as raw data. This normalized raw data is used to compute IS, IV and RA. These values are used to compute CFI. This development pipeline is shown in figure 3.

The software domain of the circadian management system is a fairly easy-to-use interface for users to just display and update the value of the final CFI that has been gathered from our trusted cloud dataset, therefore requiring just a single API call to fetch the data value alone by encapsulating the logic behind it. This ultimately avoids major flaws and enhances our time process overall.

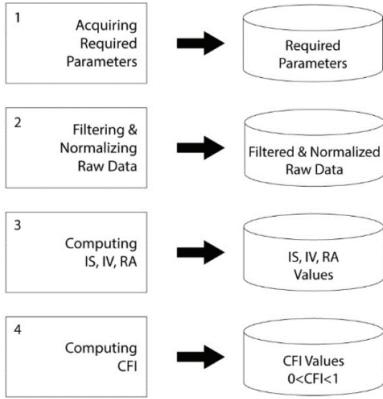


Fig. 3. Process of computing CFI value

H. IS (Interdaily Stability)

This variable refers to the stability of the activity rhythm observed over a set of consecutive days. We can plot the activity status of an individual for each day and observe the variations in the user's activity hence calculating how stable their daily routine is.

The formula for IS is shown in (1):

$$IS = \frac{n \sum_{h=1}^p (\bar{x}_h - \bar{x})^2}{p \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

Where, n: the total number of data recorded (24h * (for example)7 days = 168 data)

P: the number of data entries recorded per day (in this example, 24 as we calculate hourly)

\bar{x}_h : the mean/average of the data recorded hourly

\bar{x} : the mean of the entire dataset

x_i : the individual data value with position i.

Range: $0 \leq IS \leq 1$

I. IV(Intradaily Stability)

The activity of an individual in 24 hours can be plotted to envision the intradaily variability. For an irregular activity (which can pertain to infrequent sleep patterns throughout the day), one can notice fragmentation of the activity rhythm. These fragments will be greater in number for an individual with a high IV value.

The formula for IV is shown in (2):

$$IV = \frac{n \sum_{i=2}^n (x_i - x_{i-1})^2}{(n-1) \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

Range: $0.00 \leq IV \leq 2.00$

The IV value is normalized between 0 and 1.

J. Relative Amplitude

This variable refers to the difference between the most active 10 hours (M10) and least active 5 hours (L5) in an average lifespan of 24-hours, divided by their sum.

$$RA = \frac{M10-L5}{M10+L5} \quad (3)$$

Where, M10: average (maximum values obtained for a consecutive period of 10 hours)

L5: average (minimum values obtained for a consecutive period of 5 hours)

Range: $0 \leq RA \leq 1$

Higher relative amplitude represents better rest-activity rhythm on that day and lower relative amplitude represents worse rest-activity rhythm. Where better rest-activity represents less activity during night and high activity when awake.

K. CFI (Circadian function index)

CFI is defined as:

$$CFI = \frac{IS + RA + 1 - IV}{3} \quad (4)$$

Range: $0 \leq CFI \leq 1$

IV. RESULTS AND DISCUSSION

In our study, the participant was asked to wear the "Rhythm Monitor", a wrist-worn device for seven consecutive 24-h periods. The activity data, i.e., the number of activities per minute, was collected from the user. Thus, 7×1440 activity data, which represents 7 days of 1440 minute activity data was collected. This data was fed as input to the circadian management system which calculates the IS, IV, and RA of the data. Our circadian management system uses the ActCR package from R that has inbuilt functions for calculating IS, IV, and RA. These values are used to find the circadian function index of activity data. From the input data, the IS, IV, and RA values of 0.3228824, 1.171904 and 1 were obtained. By applying these values in (4), the CFI

$$IS = 0.3228824$$

$$RA = 1$$

$$IV = 1.171904$$

$$IV(\text{Normalized}) = 0.570952$$

$$CFI = \frac{0.3228824 + 1 + 1 - 0.570952}{3} \text{ From (4)}$$

$$CFI = 0.5839768$$

This CFI indicates that the person has an above-average circadian rhythm and the person should further improve his lifestyle to have efficient circadian rhythmicity. The above calculation is done based on the number of activities

performed by the person, while the system will calculate the CFI of all parameters that are obtained from the sensors including, temperature from the wrist temperature sensor, light intensity from the light ambient sensor, etc. Then all the CFIs calculated will be normalized into a score, which will be seen by the person.

V. CONCLUSION AND FUTURE WORKS

In this work, we presented our hardware named “Rhythm Monitor” which effectively monitors the circadian parameters such as sleep, light, wrist temperature, activity and environmental temperature. The circadian function indexes of these parameters are calculated in our circadian management system which is located in the cloud. These circadian function indexes are used to indicate the circadian robustness of the user. CFI value of zero indicates the absence of circadian rhythmicity and one indicates a robust circadian rhythm. With the proposed device, users can effectively track their circadian rhythmicity.

In addition to monitoring the CFI, a recommendation system using appropriate machine learning or reinforcement learning algorithms could be deployed on our device to further help the user make changes in their daily schedule to improve their scores.

ACKNOWLEDGMENT

This work has been funded by SRM New Gen IDEC under grant no.(NIS_3_013) to develop the prototype.

REFERENCES

- [1] National Institute of General Medical Sciences, Circadian rhythms, October 2020[Online].
Available:<https://www.nigms.nih.gov/education/fact-sheets/Pages/circadian-rhythms.aspx>
- [2] A. Martinez-Nicolas, M. J. Martinez-Madrid, P. F. Almaida-Pagan, M.-A. Bonmati-Carrion, J. A. Madrid, and M. A. Rol, “Assessing chronotypes by ambulatory circadian monitoring,” *Frontiers in Physiology*, vol. 10, 2019 [Online].
Available:<https://doi.org/10.3389/fphys.2019.01396>
- [3] E. Ortiz-Tudela, A. Martinez-Nicolas, M. Campos, M. Á. Rol, and J. A. Madrid, “A new integrated variable based on thermometry, actimetry and body position (TAP) to evaluate circadian system status in humans,” *PLoS Computational Biology*, vol. 6, no. 11, 2010 [Online].
Available:<https://doi.org/10.1371/journal.pcbi.1000996>
- [4] K. F. Swaiman, S. Ashwal, D. M. Ferriero, N. F. Schor, R. S. Finkel, A. L. Gropman, P. L. Pearl, and M. Shevell, “Actigraphy,” in *Swaiman’s Pediatric Neurology: Principles and Practice*, Sixth., Edinburgh: Elsevier, 2018, pp. 672–672.
- [5] T. Morgenthaler, C. Alessi, L. Friedman, J. Owens, V. Kapur, B. Boehlecke, T. Brown, A. Chesson, J. Coleman, T. Lee-Chiong, J. Pancer, and T. J. Swick, “Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: An update for 2007,” *Sleep*, vol. 30, no. 4, pp. 519–529, 2007 [Online].
Available: <https://doi.org/10.1093/sleep/30.4.519>
- [6] P. H. Lee and L. K. Suen, “The convergent validity of Actiwatch 2 and actigraph link accelerometers in measuring total sleeping period, wake after sleep onset, and sleep efficiency in free-living condition,” *Sleep and Breathing*, vol. 21, no. 1, pp. 209–215, 2016 [Online].
Available: <https://doi.org/10.1007/s11325-016-1406-0>
- [7] E. Ortiz-Tudela, A. Martinez-Nicolas, M. Campos, M. Á. Rol, and J. A. Madrid, “A new integrated variable based on thermometry, actimetry and body position (TAP) to evaluate circadian system status in humans,” *PLoS Computational Biology*, vol. 6, no. 11, 2010.
Available:<https://doi.org/10.1371/journal.pcbi.1000996>
- [8] J. K. Wyatt, “Delayed sleep phase syndrome: Pathophysiology and treatment options,” *Sleep*, vol. 27, no. 6, pp. 1195–1203, 2004.
Available:<https://doi.org/10.1093/sleep/27.6.1195>
- [9] M. G. Figueiro, R. Hammer, A. Bierman, and M. S. Rea, “Comparisons of three practical field devices used to measure personal light exposures and activity levels,” *Lighting Research & Technology*, vol. 45, no. 4, pp. 421–434, 2012.
Available:<https://doi.org/10.1177/1477153512450453>
- [10] A. J. Phillips, P. Vidafar, A. C. Burns, E. M. McGlashan, C. Anderson, S. M. Rajaratnam, S. W. Lockley, and S. W. Cain, “High sensitivity and interindividual variability in the response of the human circadian system to evening light,” *Proceedings of the National Academy of Sciences*, p. 201901824, 2019.
Available:<https://doi.org/10.1073/pnas.1901824116>

Identification of voice pathology from temporal and cepstral features for vowel ‘a’ low intonation

Gayathri. S
Department of Electronics and Communication
Engineering
Sri Sairam Engineering College
gayathri.ece@sairam.edu.in

Priya. E
Department of Electronics and Communication
Engineering
Sri Sairam Engineering College
priya.ece@sairam.edu.in

Abstract- In the sphere of medicine, voice pathology analysis is a very important field. Voice makes sound as it is made up of air and vibrations of vocal cords. The wind blows across the vocal cords, forcing them to come together and vibrate when a person speaks. This vibration generates the sound. Anything that causes a change in one's voice is referred to as a voice disorder. Around 25% of the world's population suffers from voice issues as a result of their jobs requiring them to talk much louder than normal. Voice assessment of vocal pathologies is an important technique for recognising voice disorders. As the voice signal has lot of unpredictability, feature extraction will lessen the variability present. In this work three different pathologies such as Laryngitis, Leukoplakia and dysphonia are considered. Pathological voice samples pronounced with sound of vowel ‘a’ uttered in /a/_1 (low pitch) is chosen from Saarbrucken voice database. Different temporal features such as signal entropy, root-mean-square, energy, zero crossing rate and Mel-Frequency Cepstral Coefficients (MFCC) are extracted from the voice signals. Results shows that among the extracted features, MFCC is able to discriminate between different voice pathologies considered.

Keywords- Voice disorders, Saarbrucken voice database, MFCC, Time-frequency audio features and cepstral features.

I. INTRODUCTION

Human are the most intelligent organisms on the planet, and one of their strongest characteristics is their ability to communicate. Speaking is the most common and easiest technique to express one's thoughts to convey an information. Every person's voice is distinct, by the way they communicate. Voice is one of the distinguishing characteristics of a person among the unique identity such as finger prints, tongue prints, and iris prints. A voice signal is a complex signal that is made up of multiple single-frequency sound waves that move together in the medium as a change in pressure. Many suffer from voice disorders from a variety of factors including severe organ damage, air pollution,

smoking, and stress [1]. Contribution to voice disorders include infections of vocal tissue, weariness, environmental changes, muscle dystrophy and face pain [2]. There is different voice disorder that occur in human voice such as Laryngitis, Funktionelle dysphonia, Rekkurrensparses, vocal fold polyp, leukoplakia, vocal nodule, Reinke edema, Spasmodic dysphonia etc. Generally, voice disorders involve issues with pitch, loudness, tone, and quality if the voice cords do not vibrate normally.

The identification or classification of voice pathologies could be spotted by the way of drawing convenient attributes. Features including temporal and cepstral features articulate relevant information from voice. Literature study reveals that majority of voice classification is based on voice features extracted. These features include signal entropy, signal energy, zero-crossing rate, Mel-Frequency Cepstral Coefficient (MFCC), perceptual linear prediction and multidimensional voice program that study diverse acoustic aspects. The goal of extracting characteristics of a voice signal improves the accuracy of voice pathology diagnosis and classification. It also evaluates the impact of different frequency regions on the detection and classification procedures. As a result, selecting and extracting voice features has become a major emphasis of voice classification [2]. Linear predictive cepstral coefficients and pitch frequency are common speech variables utilised in voice pathology detection. The MFCC replicates a human's hearing mechanism, whereas the LPCC can simulate a human's voice or speech production mechanism. Many authors have shown that MFCC is the commonly used algorithm for feature extraction of voice signal [3].

In this work voice pathologies such as Laryngitis, Leukoplakia and Dysphonia are considered for categorization. Voice signals uttering /a/_1 pitch alone is chosen from Saarbrucken voice database. Meaningful features in temporal and

cepstral are extracted and analysed for categorizing the three pathological conditions.

II. METHODOLOGY

A. Voice Database

The Saarbrucken voice database (SVD) is provided by the University of Saarland's Institute of phonetics [4]. Voice recordings of many disorders, both functional and organic, are available in SVD. Fusion of separate recordings is pronounced with normal, low, high, and l-h-l intonations. It has a collection of over 2041 people's voice recordings. It contains recordings of 1354 pathological voices from 627 males and 727 female suffering from more than 71 diseases. There are also 687 healthy voices recorded, including 259 men and 428 females. All voice recordings are sampled at a 50 kHz frequency with a 16-bit resolution. The duration of voice samples ranges from 1 to 3 seconds. The average age of the speakers is between 15-50 years old.

B. Methodology proposed

The block diagram of the proposed work is shown in figure 1. The primary step is concerned with gathering of three different pathological voice samples namely Laryngitis, Leukoplakia and dysphonia from the SVD database. The next step involves in the extraction of feature such as energy, zero crossing rate (ZCR), Root mean square (RMS), signal entropy, power and MFCC from the voice signals. The average value of each feature is computed across all data for effective categorization of pathological voice samples.

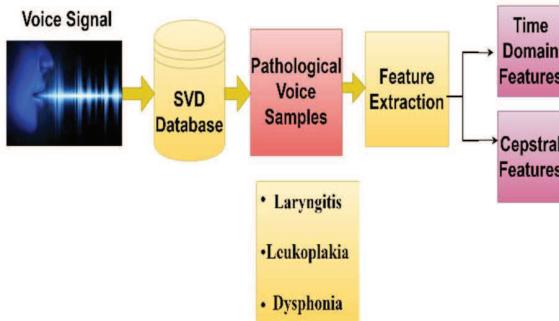


Figure 1. Block Diagram

C. Feature Extraction

1) RMS

The RMS of a signal is a measure of the signal's power content. Equation (1) represent the RMS [7].

$$RMS \triangleq \sqrt{\sum_{n=1}^N x^2(n)} \quad (1)$$

2) ZCR

The frequency content of a signal can be measured simply by the rate at which zero crossings occur is shown in equation (2). The zero-crossing rate is the number of times the amplitude of voice signals passes through zero in a given time interval or frame [8].

$$Z_n = \sum_{m=-\infty}^{\infty} |sgn[x(m)] - sgn[x(m-1)]|w(n-m) \quad (2)$$

3) SH

A signal entropy is a measurement of its spectral power distribution. The quantity of information carried by a signal is measured by its entropy [9]. Entropy is given by the equation (3),

$$\text{Entropy} = - \sum_{i=0}^{N-1} P_i * \log_2(P_i) \quad (3)$$

4) MFCC

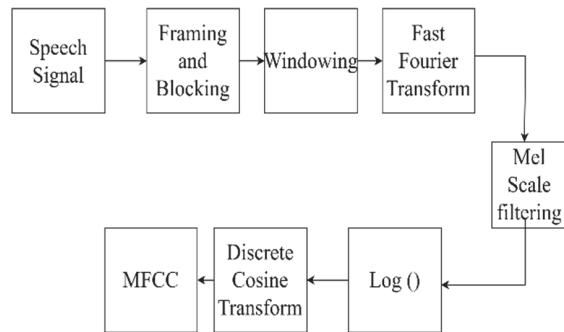


Figure 2. Mel Frequency Cepstral Coefficients.

MFCC procedure is explained in figure 2. The voice signal is derived into frames of specified window length. The voice signal is transformed by applying Discrete Fourier Transform (DFT). The DFT version is warped on a Mel scale to distinguish the higher frequency range of the audio frequency band. The inverse Fourier transform of the logarithm of magnitude spectrum is the Cepstrum. MFCC is based on varying the critical bandwidth of the human ear with linearly spaced frequency filters at low frequencies and logarithmically spaced frequency filters at high frequencies. This captures phonetically significant speech qualities [8, 11]. The equation (4) signifies the MFC coefficient representation.

$$Mel(f) = 2595 \log_{10}(1 + \frac{f}{700}) \quad (4)$$

III. RESULTS & DISCUSSION

A representation of a pathological voice signal namely Laryngitis, Leukoplakia and Dysphonia when vowel uttered /a/_1 plotted as shown in figure 3. The figure shows the time-domain representation of the three different pathological voice signal. It shows the amplitude of sound wave changing with time. Zero amplitude represents the silence observed in the voice signal.

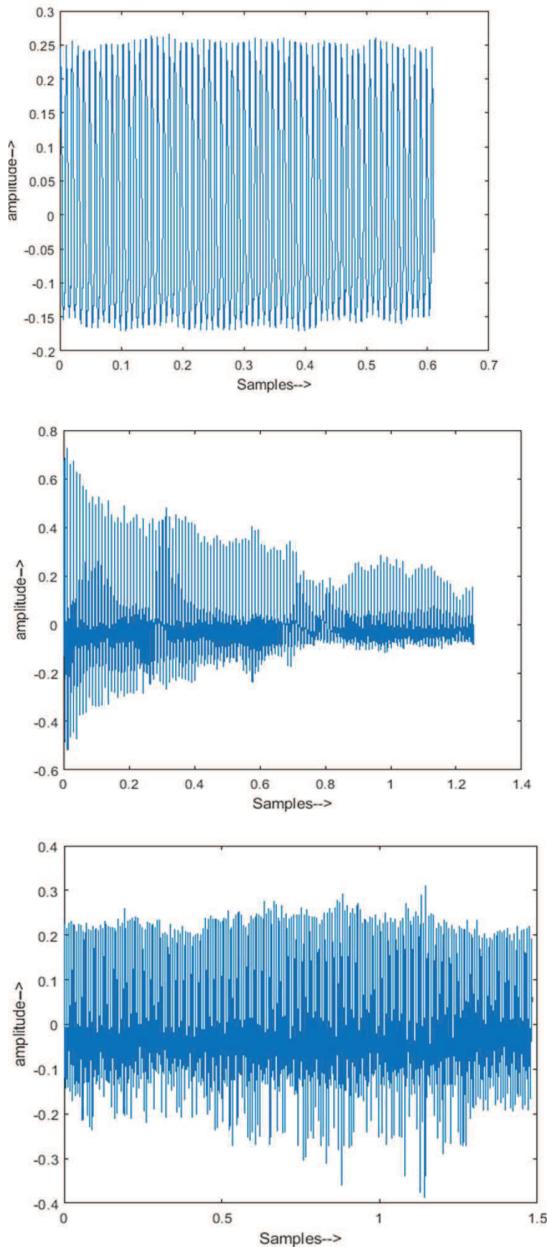


Fig.3 Voice signal for / a/_1 of (a) Laryngitis (b) Leukoplakia and (c) Dysphonia

From figure. 3 it is found that, in Laryngitis signal the vocal cord width becomes swollen and the signal amplitude varies evenly from -0.15 to 0.25. It

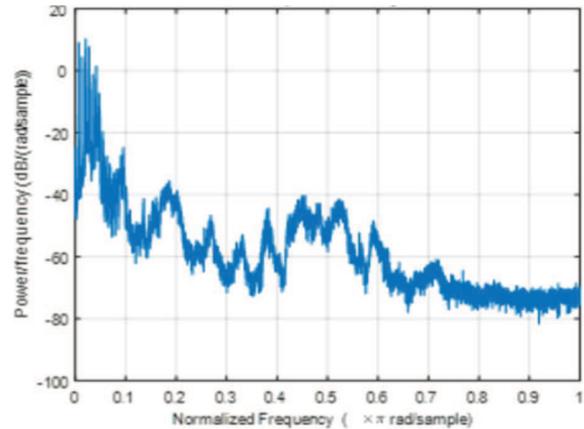
is observed for Leukoplakia, the signal amplitude varies from -0.5 to 0.65, but for Dysphonia, the amplitude varies from -0.2 to 0.2. Small jerks are observed in Dysphonia, as voice becomes quivery and jerky.

As not much inference could be derived in the time-domain representation features are derived from them. Temporal features such as ZCR, power, RMS and entropy are extracted from the voice signals. The statistical values for these features are presented in Table.1. The average value of the features power, RMS and entropy seems to differentiate pathological conditions as the mean difference between the features varies. The average value of feature ZCR is found to be same for all the three pathological samples. Hence ZCR is not a significant feature but the mean difference between the features is found to be low, thus temporal features are not much suitable for differentiating pathological voice signal.

Table 1. Average value of temporal features for three different pathologies

Temporal Features	/a/_1		
	Laryngitis	Leukoplakia	Dysphonia
ZCR	0.019±0.004	0.019±0.006	0.018±0.006
Power	0.032±0.016	0.051±0.026	0.026±0.014
RMS	2.064±0.048	0.157±0.044	0.155±0.045
Entropy	4.085±0.584	3.996±0.411	4.044±0.415

Further analysis is done by computing the Power Spectral Density (PSD) of the pathological voice signals. PSD for Laryngitis, Leukoplakia and Dysphonia are plotted in figure 4. The PSD plot shows the strength of the signals whose power is distributed in frequency domain. The power distribution is not uniform with respect to frequency for three different pathological voice signals.



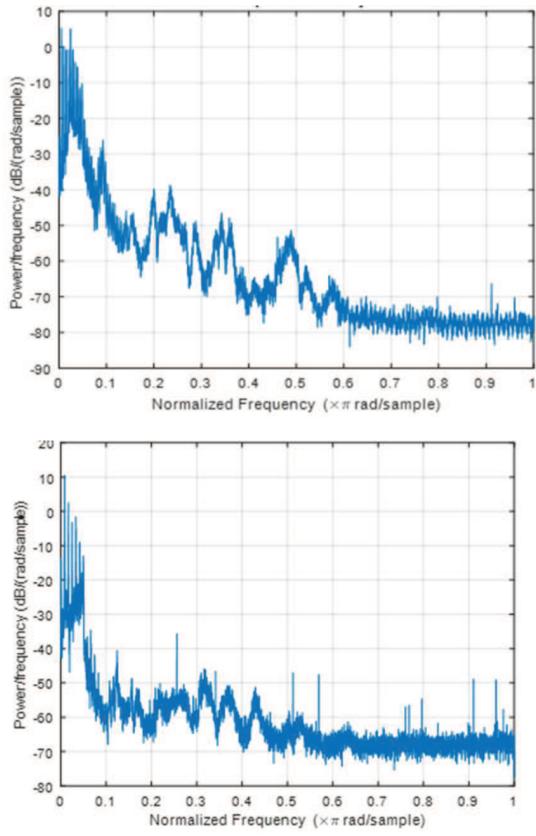


Fig.4 PSD for /a/_l of (a) Laryngitis (b) Leukoplakia and (c) Dysphonia

The PSD from figure 4 depicts the frequency spectrum of Laryngitis ranging between 0 to 0.2π radian/ sample, for Leukoplakia as 0 to 0.3π radian/ sample and for Dysphonia it ranges from 0 to 0.1π radian/ sample. This shows the pitch variation, i.e for Dysphonia the frequency is very narrow and comprises of jerky high pitch. In case of Leukoplakia, the frequency range is broad thus pitch is not pronounced for a longer time as observed from figure 3. In case of Laryngitis, pitch is observed to be in a medium level and thus time scale shows a steady amplitude. To further analyse the frequency domain variation, the time- frequency in Mel's scale is presented in Fig.5 for Laryngitis, Leukoplakia & Dysphonia.

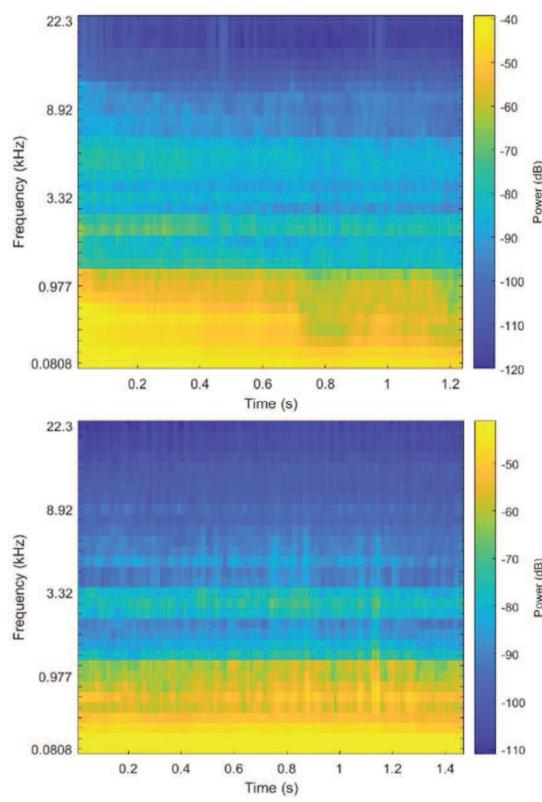
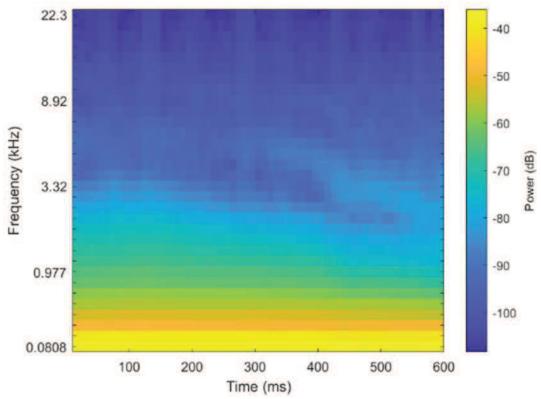


Fig.5 Mel spectrogram for /a/_l of (a) Laryngitis (b) Leukoplakiaand (c) Dysphonia

From figure 5 it is found that power in dB is shown with variation in colour of the spectrogram. The high-power signals are indicated by brighter colours such as yellow or bright orange. Low power signals are indicated by dark colours such as dark blue.

As the vocal cords become inflamed or irritated in case of Laryngitis, ranging a high-power signal is found in the low frequency from 80 to 530 Hz. Also, between 530 to 2000 Hz, spectrogram shows some visible band of energies as dark bands which again implies the infection or swelling of vocal cords. Above 2000 Hz a low power signal occurs and almost become undetectable.

In case for Leukoplakia, voice disorder is reflected as white patches or spots in the spectrogram. Therefore, figure 5 (b) shows a high-power signal at a low frequency from 80 to 1200 Hz, whereas between 1200 to 13890 Hz spectrogram shows some visible band of energies as dark bands as the vocal cords is found to be infected and above 13890 Hz the signal found to be present with lower amplitudes.

It is observed for Dysphonia in figure.5 (c), that higher power signal is found to be present in the low frequency from 80 to 1200 Hz. In case between 1200 to 13900 Hz, the spectrogram shows visible band of

energies as dark bands representing voice sounds as breathy and softer. But lower power signal occurs at high frequencies above 13900 Hz as the signal's pitch is found to be lower and hoarse.

Apart from the spectrograms, Mels cepstral coefficients of the pathological voice for Laryngitis, Leukoplakia and Dysphonia uttered in vowel /a/_1 is presented in Table.2.

Table 2. Average value of MFC Coefficients for three different pathologies.

Pathological voice and its MFCC	Laryngitis	Leukoplakia	Dysphonia
C 1	0.128 ± 0.009	0.277 ± 0.136	0.055 ± 0.021
C 2	0.040 ± 0.003	0.135 ± 0.095	0.020 ± 0.013
C 3	0.005 ± 0.0007	0.075 ± 0.070	0.007 ± 0.006
C 4	0.002 ± 0.0004	0.072 ± 0.075	0.003 ± 0.003
C 5	0.0009 ± 0.0001	0.101 ± 0.124	0.015 ± 0.014
C 6	0.0004 ± 0.0001	0.036 ± 0.047	0.020 ± 0.022
C 7	0.0002 ± 8.6058E-05	0.014 ± 0.011	0.007 ± 0.003
C 8	0.0001 ± 7.6253E-05	0.012 ± 0.010	0.003 ± 0.002

MFCC shows the information about the rate of changes in different spectrum bands. The first coefficient of MFC shows the power level of a signal from which significant difference for all the three pathological signal is found. The cepstral coefficient with a positive value, represent majority of spectral energy to be concentrated in the low frequency region. But, if cepstral coefficient has a negative value, it represents the spectral energy to be concentrated in high frequencies. Results demonstrate that lower order cepstral coefficient of C2 to C8 has a positive value indicating most of the energy to be concentrated in low-frequency region. The higher order coefficients are found not to be significant. Thus, it is observed that MFCC result presents a better demarcation in lower frequencies rather than higher frequencies.

Conclusion:

In today's world non-invasive method of diagnosis of voice pathology is an imperative tool in terms of identifying voice disorders. Audio evaluation of voice pathologies aims to reduce medical laboratory work in diagnosing pathological speeches [1]. Voice samples of three different pathologies are taken from Saarbrucken voice database, which is pronounced with vowel 'a' in low intonations. Voice samples were analysed for extraction of temporal and cepstral features energy, signal entropy, RMS, ZCR and MFCC. Results

shows that among various features for different pathologies extracted MFCC is found to be suitable for further analysis of classification process of a pathological voice signal.

REFERENCES

- [1] Al-Dhief, F.T., Baki, M.M., Latiff, N.M.A.A., Malik, N.N.A., Salim, N.S., Albader, M.A.A., Mahyuddin, N.M. and Mohammed, M.A., 2021. Voice pathology detection and classification by adopting online sequential extreme learning machine. *IEEE Access*, 9, pp.77293-77306.
- [2] Mohammed, M. A., Abdulkareem, K. H., Mostafa, S. A., Khanapi Abd Ghani, M., Maashi, M. S., Garcia-Zapirain, B., ... & Al-Dhief, F. T. (2020). Voice pathology detection and classification using convolutional neural network model. *Applied Sciences*, 10(11), 3723.
- [3] Singh, Nilu, R. A. Khan, and Raj Shree. "MFCC and prosodic feature extraction techniques: a comparative study." *International Journal of Computer Applications* 54.1 (2012).
- [4] W. J. Barry and M. Pützer. Saarbrucken voice database. Institute of Phonetics, University of Saarland. Accessed: Mar. 10, 2017. [Online]. Available: <http://www.stimmdatenbank.coli.uni-saarland.de/>
- [5] Panagiotakis, C., & Tziritas, G. (2005). A speech/music discriminator based on RMS and zero-crossings. *IEEE Transactions on multimedia*, 7(1), 155-166.
- [6] Bachu, R. G., Kopparthi, S., Adapa, B., & Barkana, B. D. (2010). Voiced/unvoiced decision for speech signals based on zero-crossing rate and energy. In *Advanced techniques in computing sciences and software engineering* (pp. 279-282). Springer, Dordrecht.
- [7] Al-Nasher, A., Muhammad, G., Alsulaiman, M., Ali, Z., Malki, K. H., Mesallam, T. A., & Ibrahim, M. F. (2017). Voice pathology detection and classification using auto-correlation and entropy features in different frequency regions. *Ieee Access*, 6, 6961-6974.
- [8] Makov, S., Minaev, A., Grinev, I., Cherhyshov, D., Kudruavcev, B., & Mladenovic, V. (2019, December). A spectral-based pitch detection method. In *AIP Conference Proceedings* (Vol. 2188, No. 1, p. 050005). AIP Publishing LLC.
- [9] Fan, L., 2020. Audio example recognition and retrieval based on geometric incremental learning support vector machine system. *IEEE Access*, 8, pp.78630-78638.
- [10] Djebbar, F., & Ayad, B. (2017). Energy and Entropy Based Features for WAV Audio Steganalysis. *J. Inf. Hiding Multim. Signal Process.*, 8(1), 168-181.
- [11] Fan, L. (2020). Audio example recognition and retrieval based on geometric incremental learning support vector machine system. *IEEE Access*, 8, 78630-78638.
- [12] Chowdhury, M. S. A., & Khan, M. F. (2019, March). Power spectral analysis and Neural network for feature extraction and recognition of speech. In *2019 International Conference on Data Science and Communication (IconDSC)* (pp. 1-7). IEEE.
- [13] Palo, H. K., Kumar, P., & Mohanty, M. N. (2017). Emotional speech recognition using optimized features. *International Journal of Research in Electronics and Computer Engineering*, 5(4), 4-9.

[14] Verde, L., De Pietro, G., & Sannino, G. (2018). Voice disorder identification by using machine learning techniques. *IEEE access*, 6, 16246-16255.

[15] Martínez, D., Lleida, E., Ortega, A., Miguel, A., & Villalba, J. (2012). Voice pathology detection on the Saarbrücken voice database with calibration and fusion of scores using multifocal toolkit. In *Advances in speech and language technologies for Iberian Languages* (pp. 99-109). Springer, Berlin, Heidelberg.

SMART VIDEO SURVILLANCE BASED WEAPON IDENTIFICATION USING YOLOV5

Dr. S. Nikkath Bushra ¹ Associate Professor Department of Information Technology St. Joseph's Institute of Technology Chennai, Tamil Nadu, India Email: ferozbushra@gmail.com	Ms. G. Shobana ² Assistant Professor Department of Computer Application Madras Christian College Chennai, Tamil Nadu, India Email: gmshobana@gmail.com	K. Uma Maheswari ³ Assistant Professor, Department of Computer Science Bharathi Women's College, Chennai, Tamil Nadu, India Email: uma.tvr1981@gmail.com	Nalini Subramanian ⁴ Associate Professor Department of Information Technolog Rajalakshmi Engineering College Chennai, Tamil Nadu, India Email id: mrgn.nalini@gmail.com
---	--	--	---

Abstract

Video Surveillance plays an important role in every aspect of life like theft detection, unusual happenings in crowded places, monitoring the suspicious activities of each individual to provide a secure and hassle free environment. Footage of closed circuit television (CCTV) camera is taken as an evidence to track the suspicious act. It is very tough to operate surveillance cameras with human intervention to detect abnormal activities. Fully automating surveillance with smart video capturing capabilities using deep learning technique is one of the most advanced means of remotely monitoring strange activities with exact location, time of event occurred along with facial recognition of criminal. Finding misdemeanor activity in a public place is very difficult to observe, as many objects are involved in the real time scenario. An uncommon or doubtful incidents in public places are captured in CCTV cameras which promotes police force to safeguard people before any mishap happens. It helps police to reach that spot on time and rescue victim. All these are meant to be achieved by using YOLO (You Only Look Once) object detection models and its variants like YOLO V1, V2, V3, V4 and latest V5 which is 88% faster than yolov4 in Deep Learning. This proposed system helps in identifying weapons held by a person as well as face recognition to identify the suspicious user. Using YOLO v5, it is very simple to track objects like weapons in a crowd. Low resolution images, far away and out of focus in the scene can also be captured and identified accurately.

Keywords: YOLO, CCTV, Weapon Detection, Real time Surveillance, Anomaly Activities.

I. INTRODUCTION

At present weapon detection at all public places includes sensors for detecting suspicious objects. Sensors are very expensive, not secure, and not efficient and also it cannot cover large area under surveillance. To overcome the drawbacks of conventional system, we focus on machine learning algorithms for object detection whose efficiency is better than using sensors alone. The algorithm is applied to all regions of input image and finds the highest score as the region of detection, which is a time consuming task for processing large number of images but with deep learning based You Only Look Once (YOLO) algorithm the object detection process is simplified by applying the algorithm to entire input image and

the region of interest is highlighted with bounding box concept to detect different real time images with higher accuracy. Through deep learning by using YOLO V5 algorithm.

Misdemeanor activities are detected easily and precisely in a Crowd. Through this algorithm high and low level objects like weapons, unusual things irrelevant to the situation is recognized and identified. This also enhances the localization tasks. In addition to this, facial recognition is also implemented through which it increases the speed by eliminating different object categories and replacing them with facial features. Thus the irregularity through webcam is monitored and is been prevented before cause. You Only Look Once (YOLO) algorithm is very popular for real time object detection. It uses YOLOv5 which is an advanced version of YOLO. YOLOv5 is fast and accurate when compared to earlier version of YOLO. The performance of YOLOv5 can be judged from Fig.1. The different variants of YOLOv5 are yolov5s, yolov5l, yolov5m, yolov5x, yolov5x.

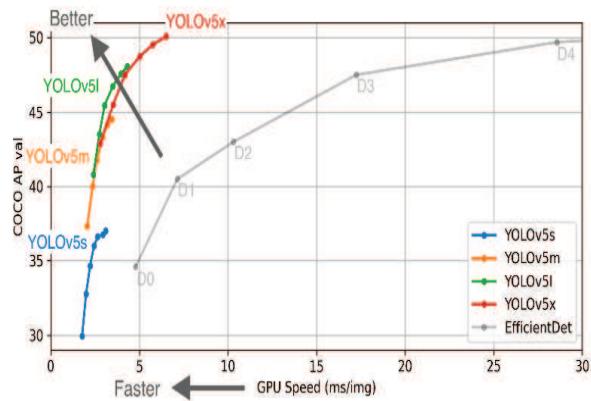


Fig 1. Performance of YOLOV5

The Difference between Yolov5 and Yolov4 is as follows
Yolov5 is very small which weighs around 27 megabytes whereas yolov4 is 244 megabytes with DarkNet architecture. Yolov5 is 90% smaller than yolov4. The overall performance can be measured by using following metrics i) mAP-mean average precision, 2) p-precision, 3) R-recall



FIG. 2 DETECTION OF OBJECTS USING YOLO

YOLOv5	YOLOv4
Size 27 Megabytes smaller	244 Megabytes(DarkNet)
180% faster(140 FPS)	50 FDPS
Accuracy 0.98 mAP	0.892 mAP

Table: 1 Yolov5 Vs Yolov4

This paper focuses particularly on three things i) weapon detection such as a person holding a knife, gun, pistol and rifles in public places ii) face detection of a person holding weapon will also be observed by surveillance camera as a suspicious even by extracting features from each segment. iii) It also monitors any suspicious activities like rising the arms suddenly, bending down, and other abnormal activities performed by a person will be discovered so it is a multiclass classification problem. Identification of specific object among several real time objects is very tough by using Video surveillance which covers multiple real time objects. This can be achieved by YOLOv5 algorithm efficiently with greater accuracy for crowd analysis.



FIG.4 DETECTING VIDEO CLIPPING OF STRANGE ACTIONS



FIG. 5 DETECTION OF KNIFE

II. LITERATURE SURVEY

The various research carried out related to this domain is highlighted in this section ,the misdemeanor activities and weapon identification along with face detection using deep learning techniques is explained.

Mohammad Tahir et al. in his research deep learning based Weapon detection especially focuses on binary classification problem like revolver or pistol as one class called pistol and another class such as wallet, metal detector, cell phone and selfie stick as non-pistol class. It uses sliding window classifiers like VGG 16, Inception V3, Inception ResNet V2 and region proposal detectors like SSD MobilNetV1, YoloV3, Faster RCNN-Inception ResNetV2, Yolo v4. Dataset used are YouTube CCTV videos, GitHub and from online movies imfdb.org. Three datasets were considered dataset 1(1732) with 750 pistol and 982 non- pistol classes, dataset 2(5254) with 3000 pistol and 2254 non-pistol, dataset 3(8327). Open CV is used for image filtering. The performance is measured in terms of recall, Mean average precision and F1 score. Yolov4 performs the best among all models with accuracy of 91.73% mean average precision, F1 Sore is 91% and confidence score of 99%. In future we can consider multi class classification problem [1].

Sri Preethaa et al. proposed automatic weapon detection in surveillance camera along with alerting system using Mobile Net architecture. Region proposal approach is used for object detection (pistol). Dataset taken from consists of 3000 images of pistol and handguns. Different models like faster RCNN, Masked RCNN, YOLO and SSD_mobilenet_v2_coco are compared and analyzed SSD_mobilenet_v2_coco as the best model in terms of speed of detection of pistol image next to YOLO model as it struggles a lot if the image is small but the overall accuracy for object classification and localization it is 95% for YOLO stands best among all other models. As future enhancement considering

environmental settings such as low resolution images with dark background [2].

H. Jain et al. present an approach for the detection of Weapon using CNN based SSD and Faster RCNN. Accuracy of 99% is achieved by faster RCNN. Dataset are taken from COCO model and also self-created images of various gun types such as AK-47 etc. In terms of speed, SSD gives better result but in terms of accuracy Faster RCNN gives 84.6 % of accuracy [3].

JunYi Lim et al. collected data from Granada dataset containing 3000 images of different guns, UCF crime dataset containing 7247 images, and own dataset with 5500 images were implemented in M2Det model under Different environmental condition were considered. Three different datasets like model 1(Granda dataset alone), model 2 (Granda with own dataset). Accuracy is good for model 2 as it contains both low resolution and high resolution images [4].

Alexander Egiazarov et al. proposed semantic segmentation model which detects a weapon especially focused on AR-15 type rifle by identifying various parts of it to easily detect the presence of object. The threshold value for weapon detection is kept as greater than 0.5 which means weapon detected else not detected. Custom dataset is created for the specific AR-15 rifle and also 4500 images from google and with the help of data augmentation generated additional 8000 samples. Negative instances is also considered which contains rifle image alone without background images. Totally 42500 image samples were taken into consideration for training network. Finally semantic segmentation and standard approach were compared for detection of each individual part of a rifle [5].

Vishwajit Dandage et al. detected presence of violence and non- violence in video surveillance by using CNN and LSTM. Faster RCNN is used for facial recognition of the person doing crime and LSTM is used as a classifier to differentiate violence or non-violence in a given video frame. If the violence is detected then an alert is sent. Own raw video frames are used as input and binary value either violent or non-violent as a predicted output value.

Kamran Ali et al. used object detection method to detect weapons like knife and gun from CCTV cameras by using various object detection algorithms along with image processing techniques like image enhancement using Wiener Filter, image segmentation with Sliding Window, feature extraction using Histogram Oriented Gradient Feature and classification using neural network to identify and locate the knife and gun from CCTV cameras by using different mathematical models. Dataset of 2400 positive training images and 1900 negative training images with 1600 test and 1200 negative test images were used.

Justin Lai et al. used overfeat network with 93% training accuracy and 89% testing accuracy. Dataset collected

from IMFDB with 3000 images for training set and 500 for validation set. VGG Net architecture using VGG-16 is utilized.

Shashank Singh Yadav et al. focused towards transportation and analyzed the performance of several road traffic irregularity detection methods based on trajectory-based abnormality detection using spatial temporal analysis using various clustering algorithms such as K-means, linear regression and Hierarchical temporal memory clustering algorithm. The object which has three-dimensional localization is considered as an incident. Unusual traffic pattern detections are framed in three different stages: Point anomaly, Sequential anomaly and occurrence anomaly [7].

Shin H.C et al. suggests that, at present video surveillance acts as a major source of detecting abnormalities taking place at various locations which gets captured/recorded instantly. It is very tough to acquire the details of each and every place without adopting this procedure. This leads to the development of a region map which captures the objects surrounding the whole surveillance zone using surveillance cameras. The algorithm to identify strange activities by learning normal conditions. The surveillance camera in each location identifies and keeps tracks of all activities in public places and creates an area object map and communicates it to the server [8]. The server combines all local maps to make a over-all map for the entire area. Different Probability Maps were extracted from global maps automatically. The activities of both normal and abnormal situations were trained.

Siddharth Shashikar et al. intends lane detection system which can keep track of objects moving in multiple lanes especially in high ways. This system gives good accuracy during day time and as well as nighttime. Though this study is a mixture of different fields in the area of vehicle detection. Presently the lane detection is handled manually, in future this kind of detection can be automated by using advanced algorithms like Hough transform [9]. At present there are no modernized surveillance cameras to capture objects so accurately. In India, the traffic surveillance is assisted with the presence of traffic police, which is a cumbersome process to manage things instantly. This results in delay in identifying victims. The existing system was unable to differentiate between the object and its shadow, in future pitfalls of this methods needs to be by implementing shadow removal techniques..

Thittaporn Ganokratana et al [10], work focusses towards capturing of objects in congested and complex environment. In training phase, they use only the frames of usual actions in order to produce their corresponding dense optical flow as temporal features. It improves the performance of anomaly localization in the pixel level evaluation by proposing the Edge Wrapping method used to reduce the noise and suppress non-related edges of abnormal objects. Additionally, our proposed method does not rely on any prior

knowledge in order to design features for the input and does not involve low-level object analysis, such as object detection and tracking.

III. PROBLEM STATEMENT

This paper focuses on detection of weapons like pistol, rifle, gun knife and sword from CCTV camera and also deals with facial recognition of a person holding weapon in a crowded place. YOLO V5 detects objects located at distant places with low quality images and as well as far apart images present in video can be identified with greater accuracy within fraction of second. The main theme behind this paper is to avoid strange activities happening in society, detecting weapons, grouping of people and threatening activities is achieved by using YOLO V5 algorithm.

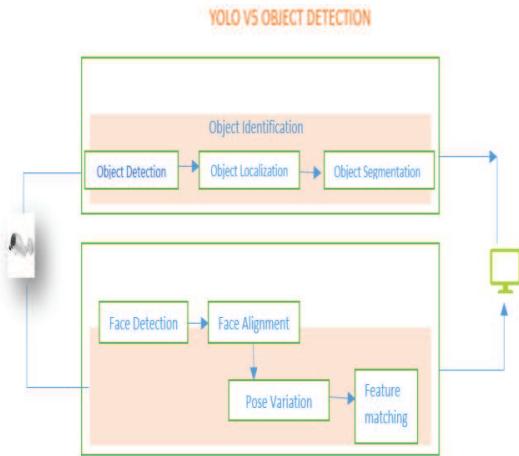


Fig. 6 Overall Yolo V5 Object Detection Model

Video surveillance for Object detection in crowd is a challenging task which needs fast and accurate real time object detection algorithms like YOLO V5 proposed by Glenn Jocher. The object detection algorithms are of two types, one is based on classification recurrent convolutional neural network (RCNN) and another is based on regression You Only look once (YOLO) which is a single shot detection algorithm, uses bounding box concept. YOLO uses Dark Net network architecture which is an open source neural network framework that supports YOLO. There are various versions of YOLO. The most recent and advanced version is YOLO V5 which is incredibly fast which can detect objects with speed of 140 fps(frames per second) and therefore fantastic tool for real-time image and video processing .YOLO V5 is considered to be the fastest tracking algorithm which is good in speed and accuracy. It is a free, open-source tool coded in PyTorch. YOLOV5 has 5 different types of models: YoloV5s(small),YoloV5 m(Medium),YoloV5 l(large) and YoloV5 x(largest).Benefits of YoloV5 is remarkable in all aspects in terms of speed, accuracy measured in terms of mean average precision (mAP)

, small in size just 27 megabytes so easily implemented in embedded devices

IV. General Working principle of YOLO

The concept behind YOLO algorithm is that it splits the input image or video into $s \times s$ grid which is responsible for detecting objects. Bounding boxes are used for object detection. The midpoint of a bounding box is responsible to detect an object which has four parameters like ‘ bx ’, ‘ by ’ represents center of bounding box,’ bw ’ is width, ‘ bh ’ is height if object is present.’ C ’ is the class object and pc is the probability of class object which tells the percentage of an object found in bounding box. The YOLO formula is

$$Y = (PC, Bx, By, Bw, Bh, C)$$

PC- Detects the presence of object in each grid its value is 1 if present else 0

Bx - Centre point in object

By - Center point in bounding box

Bw – Width of bounding box

Bh – Height of bounding box

C - Class of object either c1,c2,c3 for 3 classes
(gun, pistol and knife)

For example to detect 3 objects like gun, pistol and knife there are 3 classes for 3 objects c1(gun),c2(pistol),c3(knife). P_c is the probability of class, its value is 0 if there is no object present in the pixel. Suppose the selected pixel is

	pc
y =	
	bx
	by
	bh
	bw
	c1
	c2
	c3



Fig. 7 Gun Detection Using YOLO V5

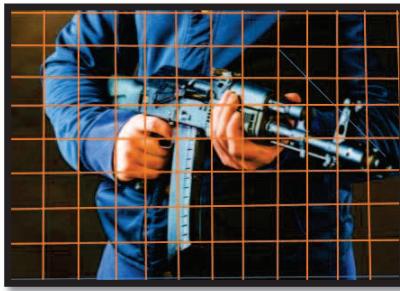


Fig.8 Bounding Box Concept

IV. DATASET DESCRIPTION:

A computer vision developer framework called Roboflow, helps in data collection, preprocessing, model training techniques and especially deploy custom dataset with ease. It has various public datasets which can be accessed by users as well as own datasets can also be uploaded. In this paper we have considered Pistol dataset of 2954 images and annotations of images from roboflow tool is copied as URL. The dataset is split into images and labels .The dataset is divided into training set and validation set. 80% of data is for training purpose and remaining 20% is used for validation. The yolov5 dataset consists of image file with its corresponding text file which has class name and bounding box coordinates. It has yaml file which stores the total number of classes to be identified and the name of each class.

V. IMPLEMENTATION DETAILS

Yolov5 for custom dataset(gun images) is implemented in pytorch. Download the whole yolov5

repository from <https://github.com/ultralytics/yolov5>. It targets at categorizing an image into one of the known number of items. It accounts to detect a way to identify specific trained objects within the given video scene. Once the module is trained with sample weapon videos it will identify those weapon objects within the current image/video depending on the value of confidence. Higher the confidence value higher the accuracy of object detected in the video.

VI. EXPERIMENTS AND RESULT

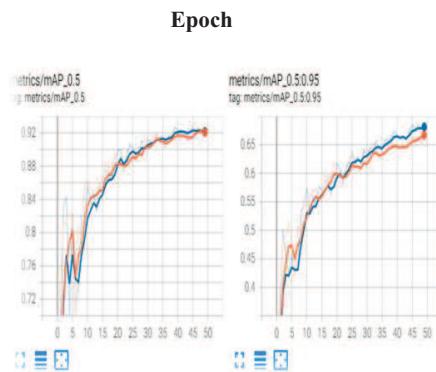
The number of epochs is set to 50 for validation of images of size 595 with 1.62 G GPU memory and image size is set to 416. The accuracy attained at mAP 0.5 is 0.98 for yolov5 and the accuracy obtained at 0.5 for yolov4 is 0.84. The size of yolov5 is 90 percent smaller than yolov4. The number of frames executed per second is 140 which is 180 percent faster than yolov4. The training loss and accuracy is obtained from tensorflow dash board.

Epoch	gpu_mem	box	obj	cls	labels	img_size
0/49	1.41G	0.08727	0.02318	0	20	416: 100% 149/149 [01:25<00:00, 1.74it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:08<00:00, 2.35it/s]
	all	595	661	0.47	0.533	0.439 0.186
Epoch	gpu_mem	box	obj	cls	labels	img_size
1/49	1.62G	0.05737	0.02176	0	16	416: 100% 149/149 [01:21<00:00, 1.83it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.57it/s]
	all	595	661	0.768	0.688	0.759 0.428
Epoch	gpu_mem	box	obj	cls	labels	img_size
2/49	1.62G	0.04962	0.01888	0	19	416: 100% 149/149 [01:21<00:00, 1.84it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.63it/s]
	all	595	661	0.82	0.773	0.837 0.52
Epoch	gpu_mem	box	obj	cls	labels	img_size
3/49	1.62G	0.04667	0.01804	0	19	416: 100% 149/149 [01:20<00:00, 1.84it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.62it/s]
	all	595	661	0.786	0.778	0.832 0.511

Fig. 9. Experimental result for 3/49th Epoch

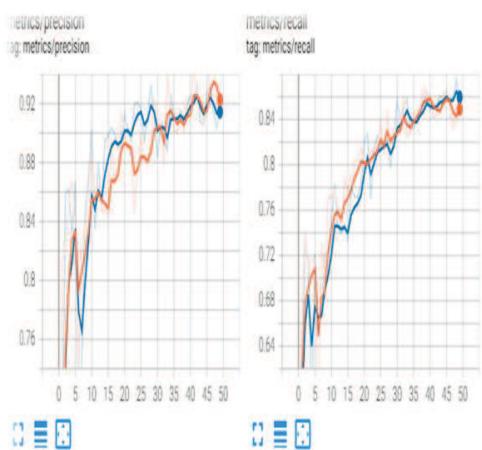
Epoch	gpu_mem	box	obj	cls	labels	img_size
16/49	1.62G	0.02311	0.01133	0	23	416: 100% 149/149 [01:19<00:00, 1.47it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.58it/s]
	all	595	661	0.903	0.864	0.924 0.683
Epoch	gpu_mem	box	obj	cls	labels	img_size
17/49	1.62G	0.02292	0.01145	0	25	416: 100% 149/149 [01:19<00:00, 1.47it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.65it/s]
	all	595	661	0.905	0.862	0.924 0.681
Epoch	gpu_mem	box	obj	cls	labels	img_size
18/49	1.62G	0.02264	0.01111	0	12	416: 100% 149/149 [01:19<00:00, 1.47it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.64it/s]
	all	595	661	0.936	0.84	0.925 0.687
Epoch	gpu_mem	box	obj	cls	labels	img_size
19/49	1.62G	0.02273	0.01124	0	15	416: 100% 149/149 [01:19<00:00, 1.47it/s]
	Class	Images	Labels	P	R	mAP@.5 mAP@.5;.95: 100% 19/19 [00:07<00:00, 2.66it/s]
	all	595	661	0.904	0.87	0.921 0.681

Fig. 10. Experimental result for 49/49th Epoch



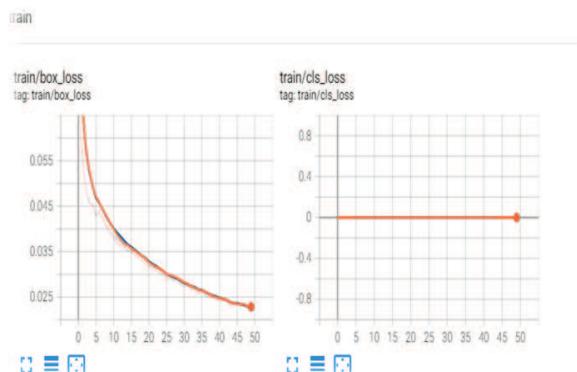
Graph. 1 Performance evaluation (mAP)

The performance of yolo v5 model is measured in terms of Mean average precision (mAP). Higher the value of mAP, more accurate the model is. The model gives 0.95 as Mean average precision.



Graph 2 . Performance evaluation in terms of Precision and Recall

The value of precision and recall varies as the number of epochs progresses. At epoch 49 precision has reached 0.98 and recall by 0.87 percent which shows a good sign of performance.

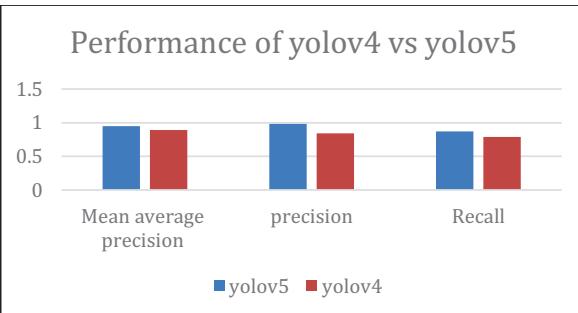


Graph 3. Training loss

VIII. CONCLUSION AND FUTURE ENHANCEMENT

By using YOLO V5 the misdemeanor activities through webcam can be identified easily on time. By using this algorithm for real-time object detection the accuracy in terms of speed has shown a tremendous improvement when

compared to all other YOLO families. A strange or unusual behavior refers to an activity that occurs in public places and gets notified immediately at all police stations and immediate action is taken to rescue from unexpected happenings, which involves sudden attack with weapons such as gun, knife and other . Identifying suspicious activities from the video frames is a challenging task can handle multiple video and image processing at the same time. This helps to achieve the misdemeanor activities happening around the environment irrelevant to the current usual scenarios. In future, low light detection can be enhanced and more advanced version of YOLO V6 can be used to achieve higher accuracy for both image and video clippings.



Graph 4. Comparison of yolov4 and yolov5

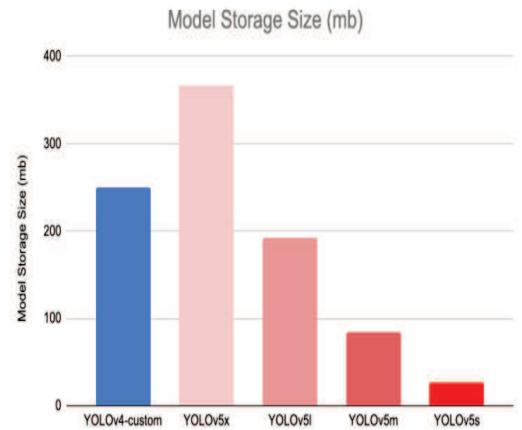


TABLE 2. PERFORMANCE OF YOLOV5 Vs YOLO V4

REFERENCES

- [1] Lianchao Jin, Fuxiao Tan, Shengming Jiang, "Generative Adversarial Network Technologies and Applications in Computer Vision", Computational Intelligence and Neuroscience, vol. 2020, Article ID 1459107, 17 pages, 2020. <https://doi.org/10.1155/2020/1459107>.
- [2] T. Nadu, T. Nadu, and T. Nadu, "Real World Anomaly Object Detection From Surveillance Video Using Tensorflow Object Detection Api," vol. XII, no. 53, pp. 53–62, 2020.
- [3] Jain, Harsh et al. "Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications." 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (2020): 193-198.
- [4] V. Dandage, H. Gautam, A. Ghavale, R. Mahore, and P. P. A. Sonewar, "Review of Violence Detection System using Deep Learning," pp. 1899–1902, 2019.
- [5] D. Zhao, J. Weng and Y. Liu, "Generating traffic scene with deep convolutional generative adversarial networks," 2017 Chinese Automation Congress (CAC), 2017, pp. 6612-6617, doi: 10.1109/CAC.2017.8243968.
- [6] H. Yousefi, Z. Azimifar and A. Nazemi, "Locally anomaly detection in crowded scenes using Locality constrained Linear Coding," 2017 Artificial Intelligence and Signal Processing Conference (AISP), 2017, pp. 205-208, doi: 10.1109/AISP.2017.8324082.
- [7] J. C. SanMiguel, J. M. Martínez and L. Caro-Campos, "Object-size invariant anomaly detection in video-surveillance," 2017 International Carnahan Conference on Security Technology (ICCST), 2017, pp. 1-6, doi: 10.1109/CCST.2017.8167826
- [8] L. Dong, Y. Zhang, C. Wen and H. Wu, "Camera anomaly detection based on morphological analysis and deep learning," 2016 IEEE International Conference on Digital Signal Processing (DSP), 2016, pp. 266-270, doi: 10.1109/ICDSP.2016.7868559.
- [9] D. Li, X. Chen and K. Huang, "Multi-attribute learning for pedestrian attribute recognition in surveillance scenarios," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 2015, pp. 111-115, doi: 10.1109/ACPR.2015.7486476.
- [10] R. Dinakaran, P. Easom, L. Zhang, A. Bouridane, R. Jiang and E. Edirisinghe, "Distant Pedestrian Detection in the Wild using Single Shot Detector with Deep Convolutional Generative Adversarial Networks," 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-7, doi: 10.1109/IJCNN.2019.8851859.
- [11] S. S. Yadav, V. Vijayakumar and J. Athanasiou, "Detection of Anomalies in Traffic Scene Surveillance," 2018 Tenth International Conference on Advanced Computing (ICoAC), 2018, pp. 286-291, doi: 10.1109/ICoAC44903.2018.8939111.
- [12] H. C. Shin, J. Chang and K. Na, "Anomaly Detection Algorithm Based on Global Object Map for Video Surveillance System," 2020 20th International Conference on Control, Automation and Systems (ICCAS), 2020, pp. 793-795, doi: 10.23919/ICCAS50221.2020.9268258.
- [13] S. S. Kalyan, V. Pratyusha, N. Nishitha and T. K. Ramesh, "Vehicle Detection Using Image Processing," 2020 IEEE International Conference for Innovation in Technology (INOCON), 2020, pp. 1-5, doi: 10.1109/INOCON50539.2020.9298188.
- [14] T. Ganokratanaa, S. Aramvith and N. Sebe, "Unsupervised Anomaly Detection and Localization Based on Deep Spatiotemporal Translation Network," in *IEEE Access*, vol. 8, pp. 50312-50329, 2020, doi: 10.1109/ACCESS.2020.2979869

Parameter Tuned Unsupervised Fuzzy Deep Learning for Clinical Data Classification

S. S. Saranya

Department of CSE, School of Computing
SRM Institute of Science and Technology
Chennai, India
saranyas6@srmist.edu.in

Dr. N. Sabiyath Fatima

Department of Computer Science and Engineering
B.S.Abdur Rahman Crescent Institute of Science and Technology
Chennai, India
sabiyyathfatima@crescent.education

Abstract—In healthcare systems, medicinal information is essential to classify and diagnose a variety of disorders at an early stage. Because of the use of Cloud-enabled Internet-of-Things (IoT) technologies, the classification of a vast amount of medicinal data has become more complex. To combat this issue, hybridized Deep Belief Network and Support Vector Machine (DBN-SVM) classifier was designed. However, its computation cost was high while initialization because the hyperparameters were not optimized. Hence this article proposes a Parameter Tuned Unsupervised Fuzzy Convolutional Neural Network (PTU-FCNN) model to classify the clinical data. First, the medical data is collected and pre-processed using a filtering scheme to obtain a clean database. Then, an Improved Principal Component Analysis (IPCA) is applied to extract the relevant attributes, which are fed to the PTU-FCNN classifier. In this PTU-FCNN classification model, a Dynamical Local-Best Harmony Search (DLB-HS) scheme is adopted to fine-tune the hyperparameters used in the U-FCNN model. This DLB-HS considers the different hyperparameters to be fine-tuned as the harmony and creates the Harmony Memory (HM) once the harmony is produced. Then, the HM is changed according to the training loss. Thus, the attributes are classified by this PTU-FCNN model to diagnose the diseases properly. At last, the testing outcomes reveal that the PTU-FCNN on the IVF dataset achieves 98.7% of accuracy than the other conventional classifier models.

Keywords—*healthcare systems, improved PCA, DBN-SVM, unsupervised fuzzy-CNN, hyperparameters, dynamical local-best harmony search*

I. INTRODUCTION

Internet-of-Things (IoT) refers to the concept of formulating and establishing web-connected devices using interactive technologies. In real-world scenarios, any individual will receive data from the other users and transfer these data to the cloud for further analysis and storage. In contrast to the classical cloud-based systems, hybrid cloud and IoT-based web applications perform well. A cloud-based IoT solution is useful for providing effective forecasting and database access from any remote location. It will be used in promising applications such as healthcare, military, and finance. Among many applications, medical data analysis is a challenging topic that must be performed effectively for illness forecasting, discovery, and diagnosis [1].

As a result, it is critical to design appropriate Artificial Intelligence (AI) models including machine and deep learning models, which accurately recognize illnesses and make a diagnosis in a timely and effective way [2-3]. Because of the increasing use of computationally

demanding approaches recently, efficient and precise processing of these health data has generated a revolution in the area of AI [4]. To aid clinicians in effectively predicting illness, a variety of diagnostic difficulties, including appropriate, robust and rapid forecasting models, must still be accomplished. The majority of healthcare databases comprise chaotic, inappropriate, overlapping, and missing data, which might affect the classifier's accuracy [5-6]. The classifier's efficiency (illness prediction) is determined by the availability of the healthcare records and the models employed in the classification phase. As a result, it is critical to employ classifiers to assess confidential clinical records in a precise and effective manner to forecast and diagnose illnesses.

To combat such problems, feature selection methods have been applied as the fundamental preprocessing stage for many AI models. These methods intend to extract and remove unwanted and overlapping attributes from the training database [7]. So, the classifier can learn the data more precisely and effectively with less computation period. Several types of research on the influence of feature selection on different classifiers reveal that feature selection not only decreases feature space but also improves classifier accuracy. To that purpose, an effective attribute mining and selection technique must be created to improve the efficiency of the resultant attribute subset. The feature selection techniques are widely separated into filter [8], wrapper [9], hybrid [10] and embedded methods [11]. Filter techniques choose relevant attributes based on the usual characteristics of the training data. Wrapper techniques create candidate attribute subsets by iteratively searching the whole attribute space. But, the filter and wrapper techniques have high computation burden, which rises as the attribute space grows. Hybrid techniques combine filter and wrapper techniques to enhance the feature selection process. Embedded techniques are those which directly choose the attributes using the optimization objective of the classifier. Such techniques achieve comparable outcomes as wrappers, yet they are less sensitive to overfitting. But, the prediction of a small set of attributes may be difficult.

Similarly, classification is a key challenge in AI models for extracting data from actual concerns. It builds a model from the data to reliably predict the target at various classes [12]. In the past centuries, many machine learning classifiers including SVM, Naive Bayes (NB), K-Nearest Neighbor (KNN), etc., have been employed to classify the healthcare records [13]. But, these algorithms have a high training time and are highly prone to overfitting while increasing the number of instances in the database.

Therefore, a boundary-based noise removal scheme [14] is presented to analyze the data and label the unknown attributes in the database. But, it has a high complexity while increasing the number of instances. So, a dual filtering technique is applied using Kalman filter and particle filtering in the preprocessing stage to remove the noises in data instances and label the unknown data. After that, an IPCA is performed to extract the relevant attributes and reduce the data dimensionality [15]. Further, the extracted features were classified by the hybrid classifier such as DBN-SVM model. But, this hybrid classification has a high computation burden during initialization since choosing appropriate hyperparameters is difficult.

Therefore in this article, the PTU-FCNN model is proposed to classify the clinical data. First, the most relevant attributes from the clinical database selected by the IPCA are fed to the U-FCNN for classification. In this classification model, the U-FCNN's hyperparameters are fine-tuned by using the DLB-HS scheme to enhance the classification efficiency. The DLB-HS scheme is a metaheuristic optimization strategy, in which the hyperparameters to be tuned are considered as the harmony and the HM is created after producing the harmony. Additionally, the HM is modified depending on the training error. Based on this optimization, the PTU-FCNN model can increase the accuracy of classifying the medical data to diagnose the diseases earlier.

The remaining sections of this manuscript are prepared as follows: Section II discusses the recent works related to the healthcare data classification. Section III describes the proposed model and Section IV illustrates its efficiency. Section V summarizes the complete work and provides the future scope.

II. LITERATURE SURVEY

Nahato et al. [16] suggested a hybrid method using Fuzzy sets and Extreme Learning Machine (FELM) to classify the medical databases. First, preprocessing was used to handle the missing values and outliers. Then, the fuzzification was performed to map all attributes to the fuzzy sets, which were further classified by the ELM. But, its efficiency was less when the number of classes was high. Xing & Bei [17] developed an improved KNN algorithm depending on the cluster denoising and density cropping to classify clinical datasets. The denoising was performed by clustering and the efficiency of KNN was increased by accelerating the search speed of KNN. But, the efficiency was influenced by the missing values in the databases.

Jiang & Li [18] designed a Radial Basis Function (RBF)-based neural network model to classify medical data. Initially, the database was preprocessed by the manifold analysis method. After that, the similarity matrix was fine-tuned using the exponential function and nearest neighbor propagation clustering. Then, the RBF neural network was built based on the variable basis width neural network structure for classification. But, the number of various types of medical data instances was analogous, whereas it needs to classify non-equilibrium medical data.

Lukmanto et al. [19] presented a classification model to detect and categorize diabetes. Initially, the diabetes database was acquired and pre-processed to solve the missing values. Then, the F-score feature selection method

was applied to choose more discriminative attributes. Further, the fuzzy-SVM algorithm was adopted to optimally learn the selected attributes and classify diabetes. But, its accuracy was not effective while using large-scale medical databases.

Zhu et al. [20] developed an improved logistic regression system to predict diabetes by combining PCA and K-means algorithms. Initially, the diabetes database was collected and pre-processed to handle the missing values. After that, the PCA was applied to remove the redundant attributes and the relevant attributes were fed to the K-means to avoid outliers. Moreover, logistic regression was used to classify the selected features to predict diabetes. But, these algorithms were suitable for a limited number of instances in the database.

Das et al. [21] designed a Linguistic Neuro-Fuzzy with Feature Extraction (LNF-FE) method to classify the medical data. At first, the linguistic fuzzification task was applied to create membership values that manage the uncertainty issues. After that, FE methods were hybridized in the NF system to choose the most relevant features, which were further classified by the Artificial Neural Network (ANN). On the other hand, the fuzzification task was not performed depending on the classes and the choice of proper membership function was difficult.

Singh and Singh [22] developed a 4-stage hybrid ensemble feature selection method to classify the clinical data. Initially, the clinical database was split by the cross-validation process. Then, different filtering algorithms were hybridized depending on the weighted scores to rank the attributes and the sequential forward selection scheme was applied as a wrapper method to get the best subgroup of attributes. Moreover, such optimal attributes were classified by the NB, SVM, random forest and KNN. But, it must choose the type of feature selection schemes among filters and wrappers to design an effective hybrid method.

Biagetti et al. [23] developed a classification model using robust-PCA feature extraction to detect Alzheimer's disorder from EEG signals. First, the medical dataset was collected and robust-PCA was applied to extract the relevant attributes. Then, different classifiers such as decision tree, KNN, SVM and NB were performed to classify Alzheimer's disorder. But, these classifiers have a high training time and computation cost while increasing the number of data. Bania & Halder [24] designed a new Rough Set Theory (RST)-based Heterogeneous Ensemble Feature Selection scheme (R-HEFS) to choose more appropriate attributes from the clinical databases. The selected attributes were then classified by the SVM, AdaBoost, random forest and NB classifiers. But, it needs an ensemble model to handle large-scale databases.

III. PROPOSED METHODOLOGY

In this section, the proposed PTU-FCNN model for medical data classification is explained briefly. First, the medical database is collected and pre-processed by the dual filtering method to solve the missing values and remove the noises in the database. Then, the cleaned database is partitioned into the training and testing sets. The training set is given to the IPCA algorithm to extract the relevant attributes from the database and reduce the data dimensionality in a given feature space.

After that, the selected attributes are passed to the PTU-FCNN classifier, which is optimized by the DLB-HS scheme to obtain the trained model. Moreover, the testing set is classified by the trained PTU-FCNN classifier to identify the type of disorder and diagnose them at an early stage. The detailed flow diagram of the aforementioned processes is depicted in Fig. 1.

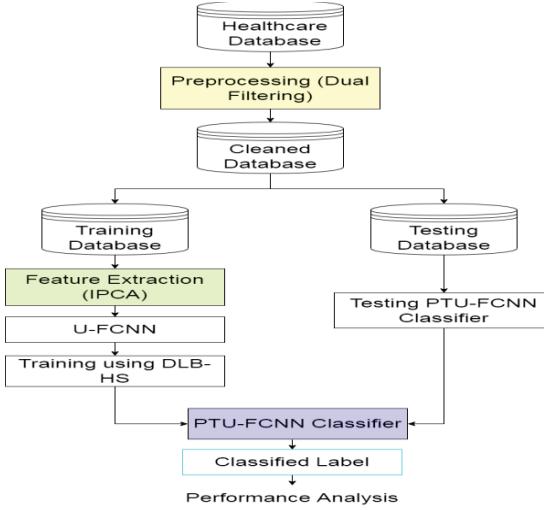


Fig. 1. Entire flow diagram of the proposed model

A. Data Preprocessing and Feature Extraction

In this study, the IVF/ICSI database [25] is considered, which comprises 100 fertility instances with 10 attributes. The attributes in this database are season (winter, spring, summer and fall), age (18-36), childish-disease (chicken pox, measles, mumps and polio), trauma (accident or severe), surgical intervention, high fever in the previous year (<3 months ago, >3 months ago and no), drinking habits (many incidences per day, each day, many incidences per week, one per week, seldom or never), nicotine addiction (never, rare or frequent), average hours sat a day (1-16) and output class (either diagnosis normal or altered). Once the database is obtained, the dual filtering using Kalman filter and particle filter is applied for discarding the irrelevant values in the database are removed to obtain the cleaned database. After obtaining the cleaned database, the IPCA algorithm is applied to extract the relevant attributes and reduce the dimensionality of the data. Further, the extracted attributes are classified by the PTU-FCNN classifier, which is described below subsection.

B. Classification using Parameter Tuned Unsupervised CNN Model

In the U-FCNN classification, the attribute maps are implanted by embedding unit to create an actual matrix. After that, the fuzzification unit is employed to convert the incoming matrix into the fuzzy space. Next, the fuzzy interpretation is convoluted with the fuzzy convolutional units and the resultant attribute maps are transformed into clear values in the defuzzification unit. Moreover, the Fully Connected (FC) unit performs as the resultant classifier of U-FCNN. The structure of U-FCNN is illustrated in Fig. 2.

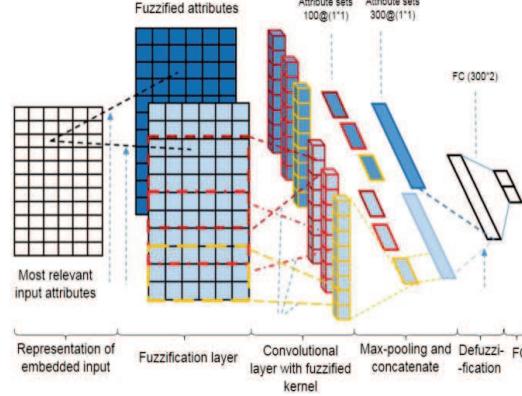


Fig. 2. U-FCNN for medical data classification

In this classifier model, the CNN comprises many hyperparameters including primary weight, training ratio, objective factor, mini-batch dimension and the amount of hidden neurons that must be determined. CNN hyperparameters need to allocate parameters like filter dimension, stride, amount of channels and zero-padding. Such hyperparameters influence the size of the attribute map. As a result, a novel method is introduced to adjust the hyperparameters in the U-FCNN using the DLB-HS scheme. The aim of this scheme is split into 3 kinds: (i) initially, the enhancement stage is achieved by basing novel enhancements on the present optimal solution vector in the HM rather than the decided memory at random, (ii) then, it splits the HM into many separately optimized sub-HMs in an effort to sustain higher range during the hunting task and (iii) it utilizes the self-training Parameter Collection Directory (PCD) to develop the HM Consideration Rate (HMCR) and Pitch Adjustment Rate (PAR) factors into the best ranges with no human configuration. The Fret Width (FW) factor is adaptively modified.

In the adaptation of the enhancement task, an arbitrary decision vector applied in the memory factor is changed using the optimal decision vector in the HM. It concentrates necessary developments on the present optimal solution, which is most certainly a local optimum and maybe the global best. To avoid potential convergence about local best, the stochastic decision vector rejected in the memory factor is rather applied in the pitch tuning. Therefore, the pitch tuning is described as:

$$i'_{new} = i_{rand} + rand() \cdot FW \quad (1)$$

In Eq. (1), i'_{new} is the new pitch-tuned range, i_{rand} is arbitrarily decided from the present ranges in the HM for a certain choice factor and $rand()$ is the random value ranging from -1 to 1. The HM is arbitrarily split into n uniformly sized sub-HMs before initializing the enhancement step. After that, all sub-HMs exploit their individual small sub-collection of decision vectors to separately join onto the local best. Because of the diversity error which occurs with constraining the size of each HMs through dividing them as the sub-HMs, it is possible that convergence cannot be towards the global best; however, instead of toward the local solution. To avoid this range of error, data is facilitated to swap among sub-HMs with a rate managed through rearranging list S .

The S measures how regularly data swap is facilitated among sub-HMs through arbitrarily rearranging the independent decision vectors of each sub-HMs into n fresh sub-HM arrangements all S iterations. It enhances the range of each sub-HM and enables the optimal decisions concurrently from the whole HM to be distributed amid sub-HMs.

After obtaining the global best, the addition of range by the rearranging process doesn't adjust; however, it may yield improper convergence. So, the DLB-HS executes the last step of the fine-tuning task after 90% of the iterations are performed. During the last step, the rearranging process is terminated and a novel HM is created by merging the optimal solution vectors from each sub-HM into the unified novel HM. After that, the novel HM is completely analyzed until the maximum iteration is achieved.

The HMCR and PAR ranges are adaptively computed through choosing from a self-training PCD. This task initiates by the initialization of the PCD, which occupies with arbitrarily created HMCR and PAR ranges. HMCR and PAR ranges are created from a regular distribution, where $HMCR \in [0.9, 1]$ and $PAR \in [0, 1]$.

At all iterations, one set of HMCR and PAR ranges are eliminated from the PCD and applied for that iteration. When the present iteration's enhancement yielded the HM to be modified, the present factor set is kept in the Winning PCD (WPCD). This procedure repeats until the PCD becomes null. After that, the PCD is refilled by arbitrarily choosing a set from the WPCD in 75% of the interval and arbitrarily creating a fresh set in 25% of the interval. To avoid HM impacts in the WPCD, the WPCD is cleared every interval the PCD is refurnished. The outcome is that the optimal parameter collection is regularly trained. Similarly, the FW parameter is adaptively tuned.

This DLB-HS helps a huge FW in the primary iterations to support the discovery of the hunt space when a diminutive FW is preferred in the last iterations to obtain superior decisions in the HM. So, the FW is regularly reduced with maximizing the iterations as:

$$FW(x) = \begin{cases} FW_{max} - \frac{FW_{max}-FW_{min}}{MI} 2x, & \text{if } x < \frac{MI}{2} \\ FW_{min}, & \text{Or else} \end{cases} \quad (2)$$

In Eq. (2), FW_{max} and FW_{min} are the highest and least ranges of the FW, x is the current iteration and MI is the highest iteration. The maximum and minimum range for the FW need to be assigned earlier than optimization begins; however, it is considered to be greatly simpler than choosing on a certain range for the FW.

Thus, by continuing the DLB-HS scheme, it is supposed that each hyperparameter kept in the HM can converge rapidly. While the iteration is terminated, the U-FCNN is designed by the optimal hyperparameters kept in the HM and trained by the selected clinical data attributes. Further, the trained PTU-FCNN classifier is applied to classify the test instances and help clinicians to diagnose the diseases appropriately.

Algorithm for Proposed Medical Data Classification Model

Input: IVF database

Output: Diagnosis normal or altered

Begin

Collect the dataset containing the fertility instances;

Apply dual filtering to obtain the cleaned dataset;

Split the dataset into training and testing sets;

for(training set)

 Perform IPCA to extract the relevant attributes;

 Feed the selected attributes to the U-FCNN;

 Optimize the CNN hyperparameters using DLB-HS scheme;

 Train the PTU-FCNN classifier;

end for

for(testing set)

 Test the trained PTU-FCNN classifier using test instances;

 Classify the label, i.e. either diagnosis normal or altered;

end for

 Evaluate the classification efficiency;

End

IV. EXPERIMENTAL RESULTS

This part displays the efficiency of the PTU-FCNN model by using the IVF database, described in Section 3.1. To evaluate its efficiency, different metrics are considered such as precision, recall, f-measure and accuracy. The comparison is conducted between PTU-FCNN and existing classifier models including U-FCNN, DBN-SVM [15], FELM [16], RBF neural network [18], Fuzzy-SVM [19], KNN [22], Random forest [22], NB [22] and SVM [22]. The considered evaluation metrics are defined below:

- **Accuracy:** It is the percentage of a proper classification of positive and negative instances to the total number of instances tested.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

True Positive (TP) and True Negative (TN) are the solutions where classifier classifies the positive and negative instances as themselves, correspondingly. False Positive (FP) is a solution where classifier improperly classifies the negative instances as positive, whereas False Negative (FN) is a solution where classifier improperly classifies the positive instances as negative.

- **Precision:** It is the percentage of properly classified positive instances to the overall expected positive instances.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

- **Recall:** It is the fraction of properly classified positive instances to the overall instances.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

- **F-measure:** It is calculated as:

$$F - \text{measure} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

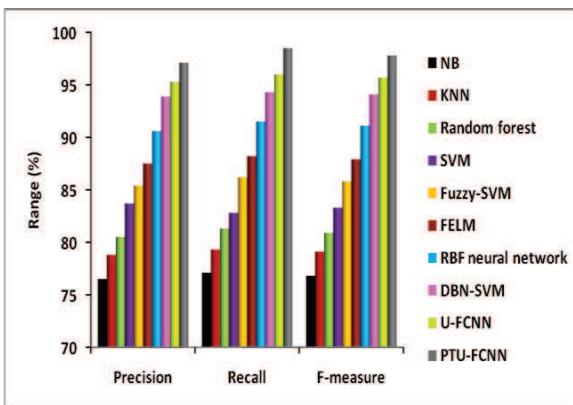


Fig. 3. Precision, recall and f-measure results for various classification models

Fig. 3 illustrates the efficiency of various models developed to classify the healthcare data. It observes that the effectiveness of the PTU-FCNN regarding the precision, recall and f-measure is superior to that of other classifier models. The precision values obtained by the NB, KNN, random forest, SVM, fuzzy-SVM, FELM, RBF neural network, DBN-SVM, U-FCNN and PTU-FCNN models on IVF database are 76.5%, 78.8%, 80.5%, 83.7%, 85.4%, 87.5%, 90.6%, 93.9%, 95.3% and 97.1%, correspondingly. Also, the recall values determined by the NB, KNN, random forest, SVM, fuzzy-SVM, FELM, RBF neural network, DBN-SVM, U-FCNN and PTU-FCNN models on IVF database are 77.1%, 79.3%, 81.3%, 82.8%, 86.2%, 88.2%, 91.5%, 94.3%, 96% and 98.5%, correspondingly. Similarly, the f-measure values achieved by the NB, KNN, random forest, SVM, fuzzy-SVM, FELM, RBF neural network, DBN-SVM, U-FCNN and PTU-FCNN models on IVF database are 76.8%, 79.1%, 80.9%, 83.3%, 85.8%, 87.9%, 91.1%, 94.1%, 95.7% and 97.8%, correspondingly.

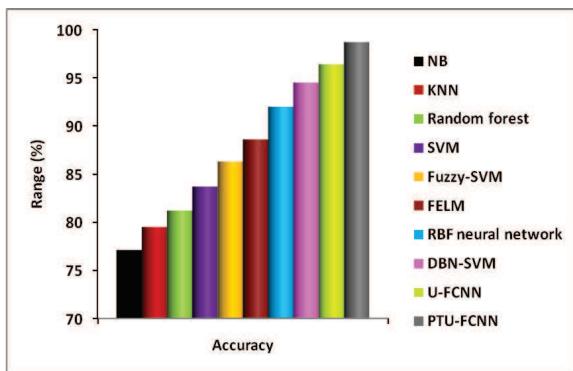


Fig. 4. Comparison of accuracy for different classification models

Fig. 4 displays the accuracy of different medical data classification models. It addresses that the accuracy of the PTU-FCNN is greater than all other conventional classifier models because of optimizing the hyperparameters efficiently. Accordingly, the accuracy of PTU-FCNN is 28.02% greater than the NB, 24.15% greater than the KNN, 21.55% greater than the random forest, 17.92% greater than the SVM, 14.37% greater than the fuzzy-

SVM, 11.4% greater than the FELM, 7.28% greater than the RBF neural network, 4.44% greater than the DBN-SVM and 2.39% greater than the U-FCNN to categorize the clinical information and treat them at prior stage.

V. CONCLUSION

In this study, the PTU-FCNN-based classification model was developed to identify and categorize healthcare information. At first, the clinical records were gathered and pre-processed to eliminate the redundant and missing data. Afterwards, the IPCA algorithm was performed on the cleaned database to capture the significant attributes. Such attributes were passed to the PTU-FCNN classifier, in which the CNN hyperparameters were fine-tuned by the DLB-HS scheme to classify the varieties of disorders and increase the classification accuracy. At last, the test outcomes proved that the PTU-FCNN model on the IVF dataset has an accuracy of 98.7% compared to the standard models for medical data classification.

REFERENCES

- [1] J. Yu, B. Fu, A. Cao, Z. He, and D. Wu, "EdgeCNN: a hybrid architecture for agile learning of healthcare data from IoT devices," in IEEE 24th Int. Conf. Parallel Distrib. Syst., 2018, pp. 852-859.
- [2] M. Nazar, M. M. Alam, E. Yafi, and M. S. Mazliham, "A systematic review of human-computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques," *IEEE Access*, vol. 9, pp. 153316-153348, 2021.
- [3] F. Shamout, T. Zhu, and D. A. Clifton, "Machine learning for clinical outcome prediction," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 116-126, 2021.
- [4] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, "Key challenges for delivering clinical impact with artificial intelligence," *BMC Med.*, vol. 17, no. 1, pp. 1-9, 2019.
- [5] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and robust machine learning for healthcare: a survey," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 156-180, 2021.
- [6] A. Vellido, "The importance of interpretability and visualization in machine learning for applications in medicine and health care," *Neural Comput. Appl.*, vol. 32, no. 24, pp. 18069-18083, 2020.
- [7] N. Cueto-López, M. T. García-Ordás, V. Dávila-Batista, V. Moreno, N. Aragónés, and R. Alazí-Rodríguez, "A comparative study on feature selection for a risk prediction model for colorectal cancer," *Comput. Methods and Programs Biomed.*, vol. 177, pp. 219-229, 2019.
- [8] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, "Benchmark for filter methods for feature selection in high-dimensional classification data," *Comput. Stat. Data Anal.*, vol. 143, pp. 1-19, 2020.
- [9] M. Kadhum, S. Manaseer, and A. L. Dalhoun, "Evaluation feature selection technique on classification by using evolutionary ELM wrapper method with features priorities," *J. Adv. Inf. Technol.*, vol. 12, no. 1, pp. 21-28, 2021.
- [10] S. Huda, J. Yearwood, H. F. Jelinek, M. M. Hassan, G. Fortino, and M. Buckland, "A hybrid feature selection with ensemble classification for imbalanced healthcare data: a case study for brain tumor diagnosis," *IEEE Access*, vol. 4, pp. 9145-9154, 2016.
- [11] S. Saito, S. Shirakawa, and Y. Akimoto, "Embedded feature selection using probabilistic model-based optimization," in *Proc. Genet. Evol. Comput. Conf. Companion*, 2018, pp. 1922-1925.
- [12] S. Gaikwad, "Study on artificial intelligence in healthcare," in *IEEE 7th Int. Conf. Adv. Comput. Commun. Syst.*, vol. 1, 2021, pp. 1165-1169.
- [13] A. Mustafa, and M. Rahimi Azghadi, "Automated machine learning for healthcare and clinical notes analysis," *Comput.*, vol. 10, no. 2, pp. 1-31, 2021.
- [14] S. S. Saranya, and N. S. Fatima, "Context aware data fusion on massive IOT data in dynamic IOT analytics," *Webology*, vol. 17, no. 2, pp. 957-970, 2020.

- [15] S. S. Saranya, and N. S. Fatima, "IoT information status using data fusion and feature extraction method," *Comput. Mater. Contin.*, vol. 70, no. 1, pp. 1857-1874, 2022.
- [16] K. B. Nahato, K. H. Nehemiah, and A. Kannan, "Hybrid approach using fuzzy sets and extreme learning machine for classifying clinical datasets," *Inform. Med. Unlocked*, vol. 2, pp. 1-11, 2016.
- [17] W. Xing, and Y. Bei, "Medical health big data classification based on KNN classification algorithm," *IEEE Access*, vol. 8, pp. 28808-28819, 2019.
- [18] C. Jiang, and Y. Li, "Health big data classification using improved radial basis function neural network and nearest neighbor propagation algorithm," *IEEE Access*, vol. 7, pp. 176782-176789, 2019.
- [19] R. B. Lukmanto, A. Nugroho, and H. Akbar, "Early detection of diabetes mellitus using feature selection and fuzzy support vector machine," *Procedia Comput. Sci.*, vol. 157, pp. 46-54, 2019.
- [20] C. Zhu, C. U. Idemudia, and W. Feng, "Improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques," *Inform. Med. Unlocked*, vol. 17, pp. 1-7, 2019.
- [21] H. Das, B. Naik, and H. S. Behera, "Medical disease analysis using neuro-fuzzy with feature extraction model for classification," *Inform. Med. Unlocked*, vol. 18, pp. 1-14, 2020.
- [22] N. Singh, and P. Singh, "A hybrid ensemble-filter wrapper feature selection approach for medical data classification," *Chemom. Intell. Lab. Syst.*, 217, 1-27, 2021.
- [23] G. Biagetti, P. Crippa, L. Falaschetti, S. Luzzi, and C. Turchetti, "Classification of Alzheimer's disease from EEG signal using robust-PCA feature extraction," *Procedia Comput. Sci.*, vol. 192, pp. 3114-3122, 2021.
- [24] R. K. Bania, and A. Halder, "R-HEFS: rough set based heterogeneous ensemble feature selection method for medical data classification," *Artif. Intell. Med.*, 114, 1-31, 2021.
- [25] DataHub. 2022. *Fertility*. [online] Available at: <<https://datahub.io/machine-learning/fertility>> [Accessed 15 January 2022].

Forecasting methods and Computational Complexity for the Sport Result Prediction

¹Vijay K
Assistant Professor
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai – 602 105, India
vijayk.btech@gmail.com

²K.Jayashree
Professor
Department of Computer Science
and Engineering
Rajalakshmi Engineering College,
Chennai-602105
k.jayashri@gmail.com

³Vijayakumar R
Assistant Professor
Department of Computer Science and
Engineering
Rajalakshmi Engineering College,
Chennai-602105
r.vijayakumar91@gmail.com

⁴Babu Rajendiran
Assistant Professor
Department of Information Technology
Rajalakshmi Engineering College,
Chennai-602105
babu.rajen17@gmail.com

Abstract— Competition in gaming is one of the most widely audience watching team game in the ecosphere. Predicting the conclusion of a contest has developed so simple because to advancements in technology. We are utilising machine learning methods such as supervised learning to envisage the winner of a One Day International (ODI) cricket match. In order to train the models, we use each player's own career statistics as well as squad recitals such as flickering and speeding performances. The ability of players and many parameters, on the other hand, play an important role in determining the ultimate outcome of a match. As a result, we're employing supervised learning algorithms to forecast the match's outcome. It also aids the team's coaches in learning and analysing where the team's performance is lacking, so that the coaches may discover a solution to increase their team's strength. So, in this study, we use four different sorts of machine learning algorithms and compare them to each other to see which one is the most accurate and produces the greatest results.

Keywords— AI and automation, Decision tree, Bayes theorem, Cricket prediction

I. INTRODUCTION

Team game in which two teams of 11 players compete against one another on a 22-yard pitch to score runs and take wickets. At a given time, the batting team has 2 players on the pitch and the bowling team has 11 players on the field one of which is a bowler. Each player on the batting team scores runs by running between the pitches or hitting boundaries(4 or 6 runs).

AI is solidly distinguished and as often as possible covers with computational experiences, which additionally centers around expectation utilizing innovation. It has strong

associations with mathematical enhancement, carries methodologies, speculation and submission regions to the field and AI is additionally partitioned into administered knowledge and unsupervised learning.

In this paper, we use parallel characterization models such as Logistic relapse, Support vector machine, Decision tree, and Bayes point machine that use directed learning calculations. A choice tree is a decision-making aid that employs a tree-like option model and its expected consequences, such as the effects of chance events, asset costs, and convenience. It is one approach to review a calculation which contains proclamations of restrictive control as it were. Choice trees are broadly utilized in activities research, particularly in choice investigation, to assist with deciding an arrangement that is probably going to accomplish an objective, yet they are likewise a well-known ML instrument. A support vector machine (SVM) is a directed prototypical of AI, which involves arrangement calculations for characterization issues in two classes. After offering an SVM model layout of named prepared information for each gathering, they're ready to categories new text. Thus, you're working on an issue with archive order. We utilize Logistic relapse to predict a parallel outcome (1/0, Yes/No, True/False) given a game plan of independent elements. Bayes point order calculation gainfully approximates the hypothetically ideal Bayesian normal of direct classifiers (with respect to theory accomplishment) by selection one "ordinary" classifier, the Bayes Point.

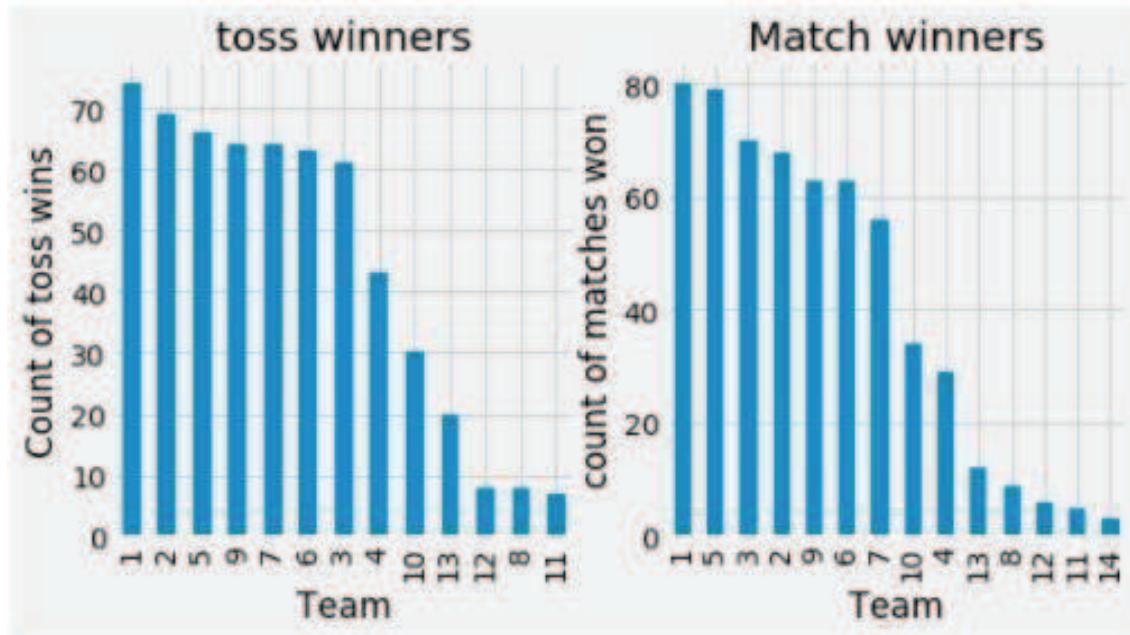


Fig. 1: Predicting the outcome of cricket matches



Fig. 2: Machine Learning

It is an information investigation device which robotizes the structure of logical models. It is a part of man-made consciousness in light of the possibility that frameworks with insignificant human information can gain from information, perceive examples and decide.

II. LITERATURE SURVEY

The Daniel [1] proposed predicting the Cricket Winner Using Machine Learning and Data Analytics, which

will aid cricketing sheets such as determining the quality of the squad and evaluating crickets.

Rameshwari Lokhande [2] proposes Live Cricket Score and Winning Prediction which briefs regarding cricket entertainment components focus on and review a few other research papers that awaited a cricket score and win.

The author [3] proposed Expanded expectation precision in the sport of cricket involving AI which endeavors to gauge player execution concerning the number

of runs every batsman can make, and the number of wickets every bowler will take for the two sides.

The author [4] proposed Prescient Analysis of IPL Match Winner utilizing Machine Learning Techniques which decides a group's previous exhibition characterizes the probability of dominating a game against a given rival.

Amal Kaluarachchi [5] proposed that they used man-made brainpower tactics to create a classification-based tool to predict the outcome in ODI cricket, principally Bayesian AI classifiers, to decide what those factors mean for the result of an ODI cricket match.

III. EXISTING SYSTEM

Cricket is a huge business market, especially in India. Hence predicting the winner of the game could be considered one of the most interesting data mining projects. There are hundreds of statistics to compare and analyse the results from, each being important in their own way. Cricket being a big money market, there is a great demand for prediction of the match winner.

Cricket data of various leagues and matches is available in different formats in various sports websites like ESPN, Yahoo Sports etc. The data is also available in a cleaned format on various other platforms. In recent times, the winner of the match is predicted either using their previous head-to-head matches or by using previous records of the matches played in that particular ground where the match is going to be held.

IV. PROPOSED SYSTEM

To predict the outcomes of ODI cricket matches, machine learning algorithms are deployed. A collection of roughly 5000 records taken from the cricinfo website. Each record contains the name of the host group, the adversary group name, throw results, the champ of the played match, edge by which the group has dominated the game, place where the match occurred, date of the match.

In this work, we basically use four attributes which provides major contribution in predicting the outcome of the match. The four fundamental and significant properties are the name of the host group, name of the rival group, match results, throws results. We utilize this information successfully and observe the result of the match utilizing regulated AI calculations, for example, Logistic relapse, Decision tree, Support vector machine, Bayes point order calculation and think about the exactness given by every one of the four calculations.

V. ARCHITECTURE

The Cricket Match Prediction and Interpretation Model Diagram employing Machine Learning as seen in Fig.3. The design consists of a database that comprises both the details and management details of the authorized users. The database also includes data from the cricket match, along with data from the squad.

The Machine Learning supervised learning algorithms such as Decision Tree, the classification algorithm of Naive Baye, Logistic Regression, and Support Vector Machine are applied to the filtered input, and the effects are evaluated. Finally, the users will render the predictions.

A. User Interface Design

The First module deals with user experience architecture for users and administrators alike. The user interface should be built in such a way that the consumer or admin will not have any difficulties in knowing what they are requested to enter and they do not check on the screen which appears to them for the things they required. Hence the user interface architecture plays a significant part in this work.

B. Adding Teams and Players

The Second Module deals with the creation of the cricket teams and the recruitment of players for the admin staff. The admin uses their user name and password to sign in to the program. When they have successfully signed in, the admin will perform two acts. The first step is that, with the addition of a team, the admin will assign teams with the necessary details (team name, country name to which the team belongs, average team winning, etc...).

The developer will always guarantee they don't add a team twice. The second step is that for the teams which they have already created, the admin will include members. A team that has at least 11 members and a maximum of 15 members must be included. The players should be included with all the required data (Player name, batting average for Batsman, bowling average for Bowler, all for All-Rounder).

C. User Registration

The third section is about the Registration of Users. Before analyzing a cricket match a person needs to have an account. To build a new user account, the user must enter the correct data (User Name, Email Address, etc...). If the user has successfully established the account, they have to login with their credentials to access the program.

D. Match Prediction

Match Prediction is dealt with in the fourth module. The root of this work begins right here. The customer must first sign in with his username and password. Upon successfully signing in, the user must pick the two teams that will be playing the cricket match. On choosing the players, the user will be shown the details of the particular match predicted.

E. Result Analysis

The final module helps the customer in evaluating the outcome of the match anticipated. Using this review the user will figure out where the success (batting and bowling) of the squad is missing. From this, they will take action to boost their team's efficiency and increasing winning chances.

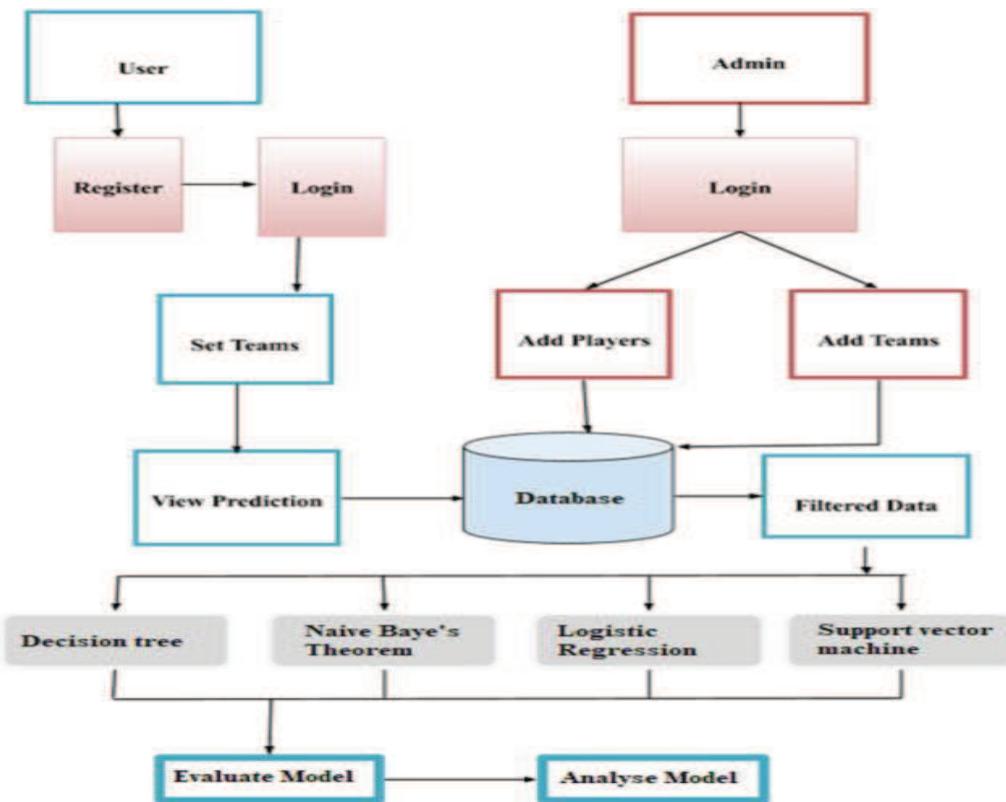


Fig. 3. Architecture of Cricket Match Prediction and Analysis

VI. MODULES

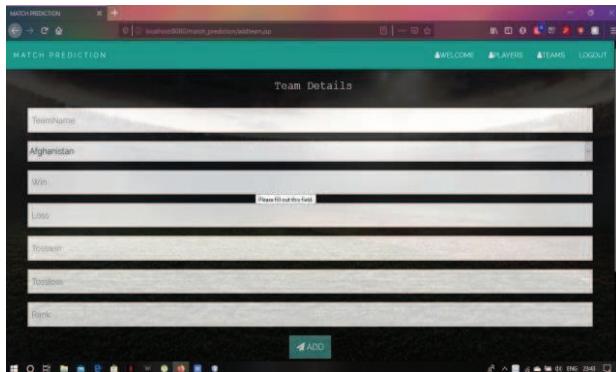


Fig. 4: Adding Teams

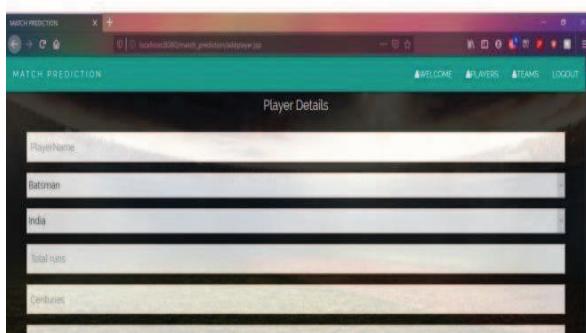


Fig. 5: Adding Players

VII. RESULT

The below Fig. 6 explains how machine learning uses the match forecast to occur in forecast and analysis of cricket match. We use four types of machine learning algorithms in this paper, and compare each other to get the best results. Firstly, the consumer will pick the cricket teams, then the outcome will display each team's winning percentage.

In table 1, On the B dataset, machine learning techniques performed better than the other three datasets. Furthermore, among all datasets, the Nave Bayes model constructed on A fared the best. For example,

the specification for first component A dataset for Accuracy is as follows:

$$A = 0.71 \times NB + 0.82 \times LMT + 0.77 \times ANN \quad (1)$$

From the Figure 7 (a) and (b), It's worth noting that making use of feature selection can help you get better outcomes. The most important features impacting match outcomes are included in Datasets A and B. This is vital information for team management as they adopt tactics to improve their players' ability.

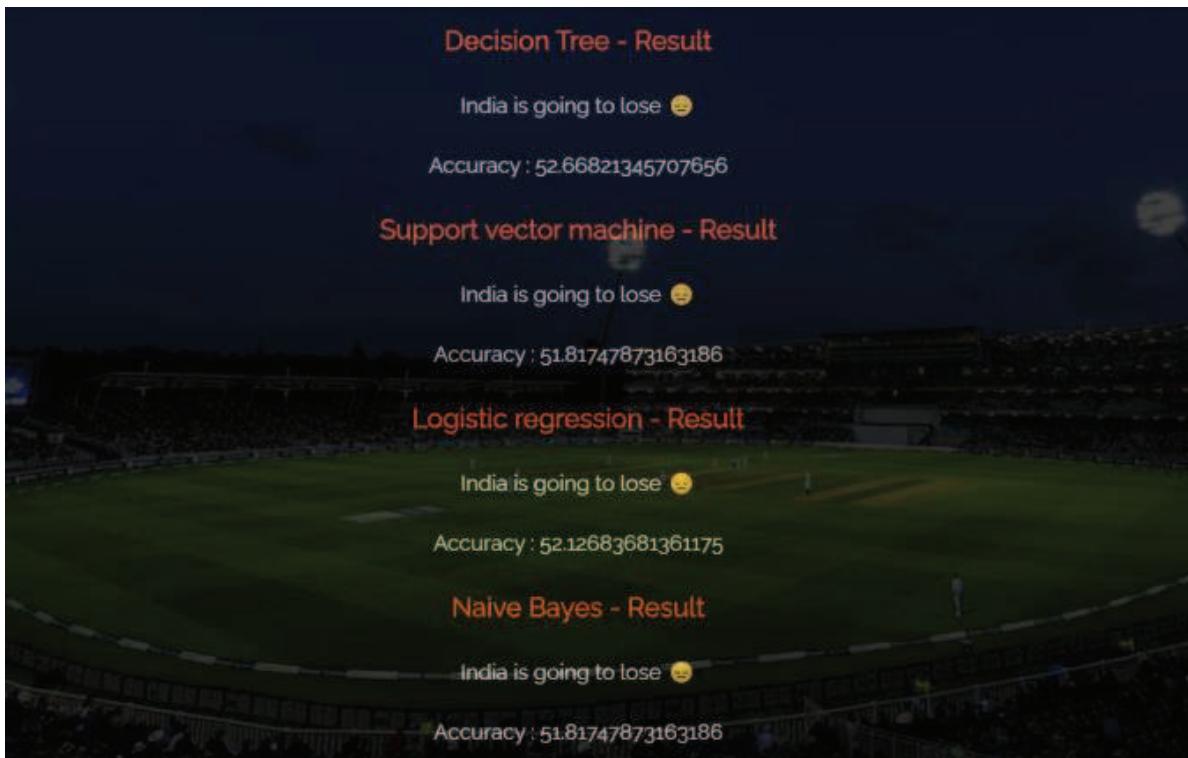


Fig. 6: Prediction Result

Table 1: The prediction accuracy

Dataset	Algorithm	Accuracy	Precision	Recall	F1
A	Naïve Bayes	0.71	0.71	0.71	0.71
	LMT	0.82	0.82	0.82	0.82
	ANN	0.77	0.77	0.77	0.77
B	Naïve Bayes	0.79	0.76	0.79	0.78
	LMT	0.80	0.80	0.80	0.80
	ANN	0.82	0.82	0.82	0.82
C	Naïve Bayes	0.73	0.73	0.73	0.73
	LMT	0.79	0.79	0.79	0.79
	ANN	0.78	0.78	0.78	0.78
D	Naïve Bayes	0.78	0.78	0.78	0.78
	LMT	0.74	0.74	0.74	0.74
	ANN	0.83	0.83	0.83	0.83

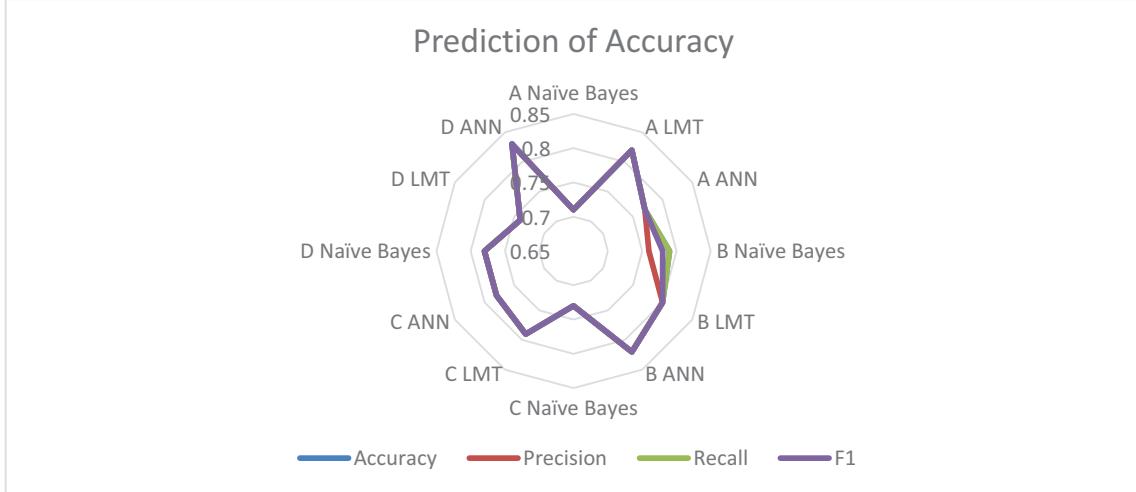


Fig. 7 (a): Prediction Accuracy

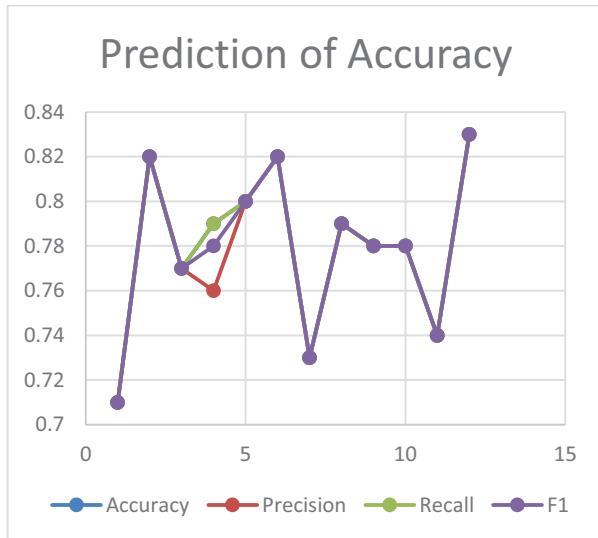


Fig. 7 (b): Prediction Accuracy

VIII. CONCLUSION AND FUTURE WORKS

This paper discusses the topic of forecasting the cricket match winner using their past success histories from match. However if they have very good track results, the players can or cannot do well in all the matches. It is one of the paper's drawbacks. This project's potential research will involve any other improved algorithms for predicting the cricket match winner with greater precision.

REFERENCES

- [1] Daniel Mago Vistro, Faizan Rasheed, Leo Gertrude David," The Cricket Winner Prediction With Application Of Machine Learning And Data Analytics",International journal of scientific and technology research volume 8,issue 09, September2019,ISSN 2277-8616.
- [2] Rameshwari Lokhande and P.M.Chawan," Live Cricket Score and Winning Prediction", International Journal of Trend in Research and Development, Volume 5(1), ISSN: 2394-9333.
- [3] Kalpdrum Passi and Niravkumar Pandey,"Increased prediction accuracy in the game of cricket using machine learning", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.8, No.2, March 2018.
- [4] Ch Sai Abhishek, Ketaki V Patil, P Yuktha, Meghana KS, MV Sudhamani," Predictive Analysis of IPL Match Winner using Machine Learning Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075, Volume-9 Issue- 2S, December 2019.
- [5] Amal Kaluarachchi, Aparna S. Varde," A Classification Based Tool to Predict the Outcome in ODI Cricket", DOI: 10.1109/ICIAFS.2010.5715668.
- [6] Kumar, S. (2014). Rankings: A batting index for limited- overs cricket.
- [7] Pardee, M. (1999). An artificial neural network approach to college football prediction and ranking.
- [8] Shah, M., 2015. Predicting ODI Cricket Result. ISSN (Paper) 2312-5187 ISSN (Online) 2312-5179 An International Peer-reviewed Journal, Volume 5.
- [9] Asare-Frempong, J. and Jayabalan, M., 2017. Predicting customer response to bank direct telemarketing campaign. In 2017 International Conference on Engineering Technology and Technopreneurship (ICE2T) (pp. 1-4). IEEE.
- [10] S. Muthuswamy and S. S. Lam, "Bowler Performance Prediction for One-day International Cricket Using Neural Networks," in Industrial Engineering Research Conference, 2008

Integrated Communal Attentive & Warning System via Cellular Systems

¹Vijayakumar R
Assistant Professor
Department of Computer Science and Engineering
Rajalakshmi Engineering College,
Chennai-602105
r.vijayakumar91@gmail.com

²Ravikumar S
Assistant Professor
Department of Computer Science and Engineering
Vel Tech Rangarajan Dr Sagunthala R &
D Institute of Science and Technology,
Chennai
ravikumars@veltech.edu.in

³Vijay K
Assistant Professor
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai – 602 105, India
vijay.k@rajalakshmi.edu.in

⁴Sivaranjani P
Assistant Professor
Department of Computer Science and Engineering
Rajalakshmi Engineering College,
Chennai-602105
ranjusiva22@gmail.com

Abstract-

In the event of an emergency, mobile text messaging systems are increasingly being used to disseminate analytical data. As a result, third-party companies claim to improve physical security by quickly delivering such communications to a variety of businesses, including colleges and universities. The primary goal of this study is to ensure user security during information sharing. The text messaging allows transmitting short, alpha numeric communication for a wide variety of applications. Public Alert System (PAS) presents two components mainly Administration Management and User Management. In PAS, there is a domain connected to intranet through which the user accounts can be created. Sending of mails to other user accounts by the authorized user is allowed and inbox can be managed. The system helps the users to view, store and delete the mails when necessary. Also, when the account is hacked, the hacker list along with the date, time of attack and the IP address is being stored in the user account. The user can view the hacker details by logging into appropriate account. In so doing, the system demonstrates that this infrastructure provides better security to the users since there is no hacking of mails.

Keywords—DNS, public alert, SMS, security, Cellular Networks

I. INTRODUCTION

The phrase "mobile computing" refers to a type of human-computer interaction in which a computer is supposed to be carried around during routine use. Mobile computing encompasses all aspects of mobile communication, hardware, and software. Mobile hardware includes device components or the mobile

devices. The characteristics and requirements of mobile apps are addressed by mobile software. The ad-hoc and infrastructural networks, as well as communication qualities and protocols, are among the communication issues. Mobility refers to the person moving between different geographical locations, networks, applications and communication devices. Also it refers to the device that moves between different networks and geographical locations. The advantage of mobile computing is mobility and it reduces the time to order for any products. Various things can be done through mobile networks such as internet, short messaging service, etc.

Computer Networks is a collection of hardware and computers in which they are interconnected by the channels of communication. It allows sharing of information and resources between different computers in the network. Communication links carry important information between machinery, control, and monitoring systems in nearly all industrial settings and facilities. The industrial settings, temperature, and liquid levels communicate a lot of control and status information. Large file transmission can also be done in the computer networks with greater efficiency.

The third-party alert system provides an efficient message delivery system during the problem of emergencies by using text messaging services. This form of communication service is offered by the cellular networks. The third-party alert system is used to send the bulk messages when there is a case of emergency like tsunami, earthquake, etc. This system provides security only in the case of smaller population and it is supposed to deliver the alert messages within ten minutes of the alert goal. If the

messages are not delivered within ten minutes, then it is retransmitted once every 15 minutes. Some of the problems in this system are

1. It does not meet the ten minutes alert goal
2. It is suitable only for smaller population
3. Congestion occurs when the messages are not delivered
4. No authentication is provided for the text messages.

Continuous values are used to represent a numeric domain. If a domain $\text{DOM}(A_j)$ is finite and unordered, it is categorical. On all categorical domains, a specific value, designated by \mathbb{Y} , is defined to represent missing values.

A logical representation of an object X is a set of attribute-value pairings.

$$[A_1 = x_1] \Lambda [A_2 = x_2] \Lambda \dots \Lambda [A_m = x_m] \quad (1)$$

where $x_j \in \text{DOM}(A_j)$ represents 1 $j \leq m$. A selector is an attribute-value pair $[A_j = x_j]$. We represent X as a vector to avoid ambiguity.

$$[x^r_1, x^r_2, \dots, x^r_p, \dots, x^c_m] \quad (2)$$

The first p items are numeric values, whereas the remaining p elements are category values.

Let $X = X_1, X_2, \dots, X_n$ represent a collection of n objects. If $x_{i,j} = x_{k,j}$ for $1 \leq j \leq m$, we write $X_i = X_k$. In the real world database, the connection $X_i = X_k$ does not imply that X_i and X_k are the same object.

The Public Alert System (PAS) is used to send alert messages to the users in an organization. It can also be used to send useful information to the users. This Public Alert System (PAS) provides better security to the users. Hacker List is used in the Public Alert System so that if any account is being hacked, an alert mail regarding hacking is saved in it. Similarly, an alert message is sent to the user's mobile number thereby providing the users, a greater security. The two major components in Public Alert System are Administration Management and User Management. Administrators have the power to identify and control the state of users logged into the system, which is known as End User.

The PAS consists of a domain connected to intranet. This domain helps in creating accounts for the organisation and users. Mails can be sent to any user within the registered companies and the inbox can be managed.

This infrastructure provides better security to the users since there is no hacking of mails. Also it does not achieve the advertised requirements for larger population. The overall determination of controls, the establishing of main objectives, and the identification of broad purposes, as well as the guidance, leadership, and control of the efforts of the groups toward some shared aims, are all the responsibility of the

administrator. Admins also have the authority to restrict users and grant them access to register their information.

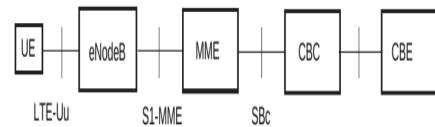


Figure: 1 Cell Broadcast network architecture in LTE. From 3GPP 23.041.
Message flow is right-to-left.

According to the User Module, a person should register their entropy, contact information, and login information. The users can operate according to the administrator's access permissions. User Management is a verification tool that allows administrators to identify and regulate the status of logged-in users.

In today's society, international communication is critical. It is possible that mail communications will take longer. It could take days, if not weeks, to get the news out to the rest of the world. E-Mail services are in charge of managing electronic communication. The technique of storing malicious programme mails in spam is included in the deduction of spam mails. Unwanted emails can be deleted to free up memory.

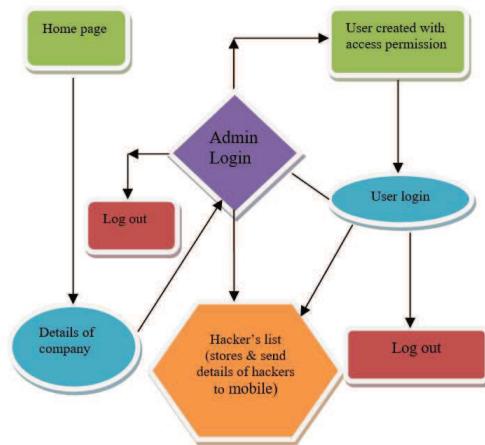


Figure 2: Flow Diagram

Acknowledgment can also be sent to pals. Incoming mail and greetings can be viewed by the user. If an exploiter's anecdote is hacked with the intent of allowing the plodder to search for the password, the user can access the hacker list details about the plodder with IP address in the relevant user account. This method is quite similar to the dictionary attack method.

The Internet Protocol address is a string of digits that is used to identify a specific machine or domain name on the Internet or Intranet. Host or network interface identification and location addressing are the two primary roles of an IP address.

II. WEBPAGE CREATION

The webpage is created for intranet mailing system. The company registration is done in this page. This company registration prompts the user to enter the company name. With the help of this company name, the user or the administrator can login into the company's home page. The next field that is to be entered is the password which requests the user to enter the similar password twice to confirm the characters.

The address of the company has to be entered for the reference. The phone number field is the next field that is prompted to the user to enter. The website address of the company is then entered. The number of employees in the company is mentioned. The category of the company has to be entered.

In order to perform email routing, DNS and the server ID is entered. Once when the above details are entered, the registration is confirmed by clicking the save button. A maximum of three companies can be created. The registered companies will be displayed at the bottom of the webpage once they are created.

At the same time, the company login is given in the home page. It prompts for the company name and password that is given while registering the company. Once when the correct company name and the password are entered, login button is clicked so that it leads to the company's web page. Forgot password can be used when the password is being forgotten and it results in the recovery of the password.

User login is also available in the web page in which the username and the password have to be entered in order to view the home page of the user. The user registration or the user creation is done by the company administration in which the details are entered by the admin of the company.

III. ACCOUNT AND MAIL MANAGEMENT

Account and mail management deals with the creation of user accounts and access permission given to the users. The company is registered and created in the home page. Created companies are viewed at the end of the home page. A particular company can be logged in so that it leads to the company's web page. The company's homepage involves 1. Creation of the user 2. Edit company information 3. Managing mails.

Creation of the user

The employees in the company can create the user id so that the employees will be able to communicate via email. It prompts to enter the username. The availability of the username has to be checked once when the desired username is entered. The password is then entered and to confirm the entered password, user has to retype the password.

The photo of the user can be browsed from the system and it can be set. The personal information of the user has to be entered. It prompts the user to enter the first name, last name and the gender. The address

for the communication is entered along with the country and the phone number. An alternate email ID is required which is meant to be optional. The security question along with the answer of that question has to be answered. This is done for the security purpose. Hacking of sms option is required. It can either be enabled or disabled according to the user's desirability. The registration of the user is now complete. If the hacking sms option is enabled during the user creation, only then the message is sent to the user's mobile number during the hacking attacks. Similarly, camera option is also provided to give better security for the users during login process.

After the completion of the creation of the user or the employee, the employee list that is created is displayed in the company's home page. Multiple employees can be created for a particular company and the list is displayed. Employee rights can be given to a particular employee. The employee is selected. The username and the password of the employee are mentioned. Compose rights, inbox rights and sent item rights are available. Those rights to be given to a particular employee can be made enabled so that the employee enjoys all the rights that are given by the administrator. The above rights mean that the employee can compose the mail, send the mail to any user, and the sent mail is then saved in the sent item. The employee can modify the contents of the inbox and the sent item if the right is provided to the employee.

Edit Company Information

In this, the information of the company is edited. The username of the company can be edited along with the personal details that are entered during the registration or the creation of the company. In the user's home page, the user can compose the message and send to the desired user. The DNS and the server ID can be edited in the company information. The number of employees can also be edited. The account can also be deleted while editing the information of the company. This can be done by clicking edit company information.

In the user's home page, the account information and the personal information of a particular employee can be edited. While editing the account information the username along with the password can be changed. Also, the security question that is given during the registration can be edited.

The personal information of the employee can also be edited. The first name and the last name of the user can be changed. Gender, address, phone number and the alternate email option can also be edited from the details that are entered during the creation of the account of an employee.

While composing the message in the 'to' field it is necessary to mention to whom the message has to be sent and in the subject field the description about the mail can be given.

Also there is field as ‘bcc’ in which another user’s mail ID can be mentioned so that the mail will be sent to the user that is mentioned in the bcc field. Hence the mail can be composed and sent to the multiple users.

In the same way, the sent mail can be viewed with the help of sent items in the account of the user. Once the mail is sent to any user, then immediately the mail will be saved in the sent item along with the details that to whom the mail is sent at what time and the date is also mentioned.

Once the mail is sent by the sender from his/her mail then the receiver will be able to view the mail that is sent by the sender. The receiver can then compose message if the employee right is applicable to the particular user.

Managing mails

E-Mail is the process of sending digital messages from a sender to one or more receivers. Postal communication may take more time. It may be completed in days or weeks to make the message available to others. But E-Mail services take lesser time to deliver the message to the recipients.

User’s mail box is fully controlled by the administrator. The user can compose mail if the permission is given by the admin. Mail communication between the users of the registered companies is allowed. Deduction of spam mails module filter the mail that comes to the inbox. If the mail is found to be a malicious program, then the particular mail is stored in the spam. Deletion of mails is allowed to manage memory.

User folder is an option provided to all the users. This option acts as a label and the users can move the important mails to this user folder.

IV. MOBILE INDICATION HACKER LIST AND IP ADDRESS TRACKER

Mobile numbers which are entered during the registration of the user is used to send the text message when the user account is hacked. And also, hacking sms option should be turned on during registration to receive this alert message. The hacker may hack the user account by trying different possibilities. This kind of attack is said to be dictionary attack.

When an account is attempted for hacking, the hacker list is stored in the respective user account. The hacker list comprises of user name, password tried, date and time of attack and the IP address of the system from where the attack is made.

The company and the employees for each company are created. The company can be logged in with the particular company name and the password. Similarly, the user can log in with the particular username and the password created by the administrator.

Table 1: Assessment of ground-truth information cast-off for calculation and working out in the literature.

Procedure	Area Measured	Training Scopes	Endorsement Scope	Testing Scope
GeoPing	US	240	-	181,246†
ShortestPing	US	-	-	8426
Path-Latency	Europe	-	-	1245
Constraint-Based	US+WesternEurope	124	-	126
Topology-Based	Europe	-	-	8541
Octant	NorthAmerica	114	-	95
Street-Level	US	-	-	125
StatisticalGeo location	US	95	-	65
Learning-Based	US	2854	-	13,350*
Posit	US	96	-	418

†-data was sourced from website user data, including that of Hotmail.

*-ground-truth data was derived from MaxMind lookups on IP addresses.

Table 2: CENAPRED EWS

Classification	Phenomenon	Data	Exposure	Twitch Time	Period of Announcement
Countrywide Seismological Service	Seismic	www.ssn.unam.mx	Countrywide	1910	a forewarning of an event
Mexican Seismic Warning System (SASMEX)	Seismic	www.cires.org.mx	Capitals of Mexico, Oaxaca, Chilpancingo, Acapulco and Morelia	1991	A few seconds before an earthquake that has already happened. It is determined by the epicenter's distance and the magnitude of the earthquake.
Popocatepetl Volcano Monitoring System	Volcanic	www.cenapred.gob.mx	The Volcano's Surrounding Areas	1994	a forewarning of an event
SIAT-CT	Humid cyclone	http://smn.cna.gob.mx www.cenapred.gob.mx	Countrywide	2000	72-hour notice is required.
National Tsunami Warning System	Tidal wave	www.bit.ly/1w3MNJa	Mexican Pacific Coast	2013	Minutes in advance for local tsunamis; hours in advance for regional, distant, or transoceanic tsunamis.
Early Warning System for fires in Mexico	Jungle fires	www.conabio.gob.mx	Countrywide	1999	a forewarning of an event
Meteorological National Service	Meteorological	http://smn.cna.gob.mx	Countrywide	1877	Prognoses and warnings prior to the occurrence

Source: CENAPRED¹

¹ In <https://www.gob.mx/cms/uploads/attachment/file/111703/298INFOGRAFASISTEMASDEALERTATEMPRANA.PDF>

Table 3: Procedural Classification

Procedures	Toolkits
Reconnaissance	Digital scanner, Social business, Dumpster diving
Probe	Scanners, Sniffers
Toehold	Spoofing tools, Malicious applets and scripts, Buffer overflow tools, Password crackers, Software bugs, Trojan horses, Holes in trust management
Advancement	Password crackers, Software bugs
Stealth	Stealth and backdoor tools
Listening post	Stealth and backdoor tools, Sniffers and snoopers, Trojan horses
Takeover	Scanners, sniffers, spoofing tools, malicious applets, buffer overflow tools, password crackers, and other similar tools are available.

Table 4: Outcomes of analysing IPs in netblocks.

Web Server IPs			Result			
Nation Cryptograph	Total	Recept ive	Unfi tting	Questi onable	Anycast	Precise
KY (Cayman Islands)	405	308	175	98	1	34
SC (Seychelles)	3354	2205	2099	94	11	1
VG (Virgin Islands)	1365	1247	1196	49	1	1
DM(Dominica)	1056	1019	998	17	2	2
Total	6180	4779	4468	258	15	38

While logging into the company login or into the user login, there might be chances of entering the wrong password. Also there may be possibilities of a wrong user to login with a wrong password. Once when the person tries to enter a wrong password with the username, then an indication will be sent to the mobile. At the same time the hacker list will be created in the user account.

The user identifies that the account is being hacked with the message that is sent to the mobile. Also, when the user login the account with the correct username and the password, there will be an option as hacker list in the homepage of the user. If the user clicks the hacker list, then the list of hackers will be displayed.

In that, the date and time of hacking will be shown. The passwords tried by the hacker will also be displayed.

$$\sum_{i=1}^n d1(X_i, Q) = \sum_{i=1}^n \sum_{j=1}^m \delta(x_{i,j}, q_j) \quad (3)$$

$$= \sum_{j=1}^m \sum_{i=1}^n \delta(x_{i,j}, q_j) \quad (4)$$

$$= \sum_{i=1}^m n \left(1 - \frac{n_{q_j}}{n} \right) \quad (5)$$

$$= \sum_{i=1}^m n (1 - f_r(A_j = q_j | X)) \quad (6)$$

With the hacker list displayed the user can track the IP address from which the particular account is hacked and will be able to find the hacker. This provides a greater security to the users from hacking.

The National Center for Catastrophe Prevention and Mitigation is the Mexican entity in charge of disaster prevention and mitigation (CENAPRED). Its mission is to support the National Civil Protection System's technical requirements (SINAPROC). Early warning systems rely on CENAPRED, as seen in Table 2.

A hacker toolkit can be utilised in one or more penetration phases, as illustrated in Table 3, and different penetration steps normally require distinct hacking toolkits. An intruder wants to obtain information about the target system or network during the reconnaissance phase. He'll need scanners to gather data from the target's machines, user accounts, and services. He may also use social engineering and dumpster diving to aid in the data collection process.

Then, in the second step, he looks for flaws in the system. He captures the actions of the target system and network with scanners and sniffers, then analyses probable security flaws and vulnerabilities.

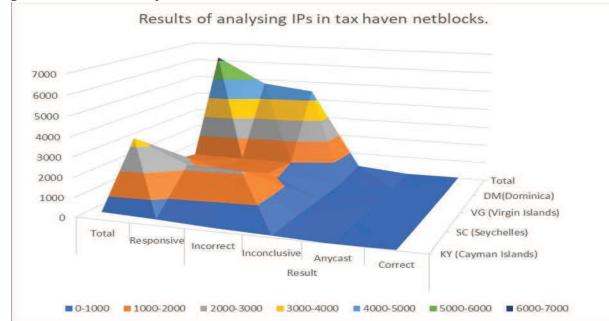


Figure 3: Outcomes of analysing IPs in netblocks.

The findings of this investigation are shown in Table 4 and Figure 3. 74.7 percent of web servers responding to ping measurements in these countries utilised clearly incorrect geolocation. With a rate of 98 percent, the Virgin Islands recorded the highest proportion of manifestly incorrect geolocation. Two of the largest blocks of IPs with incorrect geolocation are 165.231.0.0/16 and 196.196.0.0/16. Both of these netblocks cover a total of 131,072 IP addresses and have Seychelles country codes. Every IP in these ranges that serves web content is geolocated incorrectly. FiberGrid, a Seychelles-based organisation, owns both of them. The addresses appear to be in Sweden, as evidenced by the geolocation result shown in Figure 3.

VI. CONCLUSION

It is concluded that the system is functional and meets the needs of the users. In the intranet, the system has been thoroughly tested. Hackers can bypass the web server and go straight to the database server. It is anticipated that the assaults will not be blocked or detected by the present web server IDS, that the hacker will seize control of the web server after the attack, and that full control of the web server will be obtained afterwards to launch additional attacks. More than one system accesses the site at the same time. Simultaneous intranet login is investigated.

The site operates in accordance with the restrictions set forth in their respective browsers. In the future, improvements to the user's application may be made so that the website performs in a more appealing and useful manner than it does now. The transactions are now moving at a faster pace. Security can also be provided based on the needs of the users. It also aids us in locating the hacker in intrusions involving IP addresses.

REFERENCES

- [1] D. Andersen, "Mayday: Distributed Filtering for Internet Services," Proc. USENIX Symp. Internet Technologies and Systems (USITS), 2003.
- [2] T. Anderson, T. Roscoe, and D. Wetherall, "Preventing Internet Denial of Service with
- Capabilities," Proc. ACM Workshop Hot Topics in Networking (HotNets), 2003.
- [3] K. Argyraki and D.R. Cheriton, "Scalable Network-Layer Defense against Internet Bandwidth-Flooding Attacks," ACM/IEEE Trans. Networking, vol. 17, no. 4, pp. 1284-1297, Aug. 2009.
- [4] A. Stavrou, D.L. Cook, W.G. Morein, A.D. Keromytis, V. Misra, and D. Rubenstein, "WebSOS: An Overlay-Based System for Protecting Web Servers from Denial of Service Attacks," J. Computer Networks, Special Issue on Web and Network Security, vol. 48, no. 5, pp. 781-807, 2005.
- [5] A. Stavrou and A. Keromytis, "Countering DOS Attacks with Stateless Multipath Overlays," Proc. ACM Conf. Computer and Comm. Security (CCS), 2005.
- [6] P. Traynor, W. Enck, P. McDaniel, and T. La Porta, "Exploiting Open Functionality in SMS-Capable Cellular Networks," J. Computer Security, vol. 16, no. 6, pp. 713-742, 2008.
- [7] P. Traynor, W. Enck, P. McDaniel, and T. La Porta, "Mitigating Attacks on Open Functionality in SMS-Capable Cellular Networks," IEEE/ACM Trans. Networking, vol. 17, no. 1, pp. 40-53, Feb. 2009.
- [8] TXTLaunchPad, "TXTLaunchPad Provides Bulk SMS Text Message Alerts," <http://www.txtlaunchpad.com>, 2007.

REAL-TIME SMART ATTENDANCE MONITORING SYSTEM WITH THERMAL SCANNING

Gorla SujanSouri, Gummala Sainath Reddy, Panditi Prem Sagar, Dr. V. Hima Deepthi
Electronics and Communication Engineering Department,
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India.
Email: 9918005133@klu.ac.in

ABSTRACT:

In conventional attendance system, taking roll call in class time or monitoring the presence of student/employee in an educational institute or an industry has been a time-consuming process. It is also very important to see that all educational institutes are concerned about the consistency of student's performance. The main cause of this is decrease is due to inadequate attendance. So, an advanced student attendance checking system is required which supports the faculty to keep records of attendance. In this project we have introduced an intelligent face recognition-based attendance system. We have proposed to implement a "Smart Attendance through Face Recognition" which uses viola-jones algorithm for training and detection of faces. Because viola jones is highly strong, and its use has proved to be particularly noteworthy in real-time detection. The present implementation includes time-saving facial identification and eradicates the possibilities of proxy attendance due to facial authorization. In this pandemic season to alert the staff about the infected covid patients we are using thermal scanning to detect the temperature of a person. So that we would be the reason to decrease the spread of covid and save innocent lives. This system can now be used in an area in which participation plays an important role.

1. INTRODUCTION

The accuracy of the data collected was the major concern with the prior attendance tracking system. This is because attendance may not be recorded personally by the original person; in other words, a third party may take a particular person's attendance without the institution's knowledge. This compromises the data's accuracy. Figure 1 shows the pie chart depicts the results of our survey taken from people from various fields. From the above results shows that existing

fingerprint and face recognition-based systems are not appropriate and efficient in real world. The reason that majority of the people opted automatic attendance system is because of its convenience and time reducing ability. The primary goal of this study is to provide a replacement for the traditional attendance tracking system.

In the learning and teaching environment, an automated student attendance system is required. The proposed system's purpose is to address the concerns by introducing facial recognition into the attendance management process, which can be used to save time and effort during tests or lectures.

As a result, the suggested approach will assist the current student attendance system in the below points:

- Reducing the amount of time spent on attendance tracking while increasing the amount of time spent on real teaching.
- Improving security so that attendance can be tracked without the kids' knowledge.

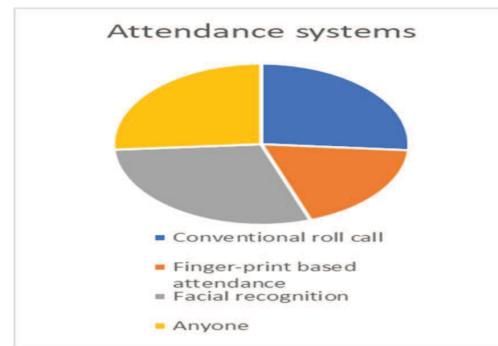


Fig.1. Pie chart showing preferences of attendance system

2. LITERATURE SURVEY

Louis Mothwa and Jules-Raymond Tapamo et al monitoring student attendance is still an important and necessary aspect of any educational institution.[1] As feature extraction approaches, PCA, LDA, LBP, and a combination of PCA and LDA were investigated. Ninety percent of the items were recognised by the model.

Payal Patil and S. Shinde et al explains Attendance reporting is a common practice at academic institutions all around the world. The manual attendance system, which is time-consuming and difficult, is commonly utilized in the school system [2]. When compared to other biometric systems, the automatic attendance system's fundamental premise is to use facial recognition with ease. Face detection accuracy is evaluated using three methods: Convolution Neural Networks, Viola-Jones (Haar Cascade), and Histogram Oriented Gradients (HOG) (CNN).

S. Matilda and K. Shahin et al describes that this technology provides a cost-effective and effective technique of recording attendance. Face recognition is the process of recognising people based on their facial features [3]. It works with the face's most obvious individual identity. The student attendance observation system is a computer-assisted tool for identifying helpful students in a complex picture or video frame. The image of a group of students is taken, and they are also identified on an individual basis. We tend to train as soon as the faces recognised fit the training information set.

Manh Cuong Le and My-Ha Le et al proposes based on data from an above camera, an autonomous people recognition and counting technique. The purpose of this research is to develop a smart people counting approach that can be used in office and lecture room attendance tracking systems [4]. The detecting step was carried out using MobileNetv2-SSD, a deep learning architecture. If a person is spotted, the tracking phase is initiated, which uses visual-tracking algorithms to maintain track of the people's positions.

In this paper, Syam Kakarla, Priyaranjan Gangula and M.Sai Rahul et al describes an IoT-based Smart Attendance Monitoring and Counting System [5]. A counter system is used to count objects in a classroom, theatre, office, mall, or

industry, among other places. In today's world, whether a facility is being used as a mall, a classroom, or a factory, proper power management is necessary. Anyone can remotely monitor the presence of people on the premises using the designed system and the Internet of Things. This mechanism determines the person's entrance and exit.

3. METHODOLOGY

To recognise and detect faces, the Viola Jones algorithm (HAAR-Cascade) is employed. The HAAR cascade is a machine learning technique that involves training a cascade function. The entire process is shown in figure 3.1. On all the training photos, we use every feature of the algorithm. At the outset, each image is assigned the same weight. It determines the appropriate threshold for categorizing the faces as positive or negative.

Both the trainer and the detector are provided by OpenCV. The cascade image classifier has two primary states. The first is training, and the second is detection. OpenCV includes two OpenCV Haart training and OpenCV train cascade apps for training cascade classifiers. The classifier is saved in a distinct file format in these two applications.

There are two different kinds of samples:

- Negative samples: non-object images are the subject of the negative sample.
- Positive samples: It's an image containing detectable items that's related

3.1. TYPES OF FEATURES AND THEIR EVALUATION

The Viola–Jones technique has the following characteristics that make it a good detection algorithm:

- Robust
- Real-time

Face detection alone (not recognition).

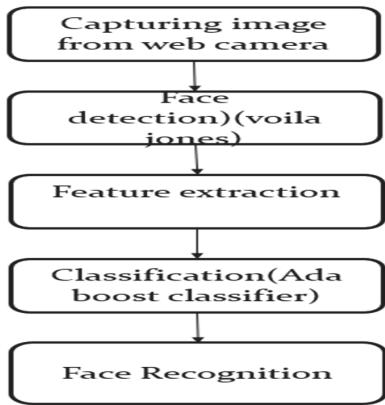


Fig.2. Block diagram of Automatic face recognition

There are four phases to the algorithm:

3.1.1. HAAR FEATURE SELECTION

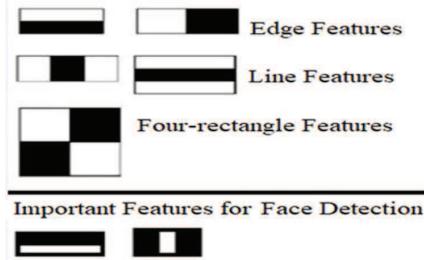


Fig.3. HAAR Features

- All human faces have some characteristics in common. As demonstrated in figure 3, these commonalities may be replicated using Haar Features.
- Human faces share a few characteristics:
- Size and location: eyes, mouth, nose bridge
- Value: pixel intensity gradients that are directed.
- The four characteristics that this algorithm matches are then looked for in a facial picture (shown at right).
- Rectangle characteristics include:

3.1.2 INTEGRAL IMAGE

- To make an Integral Image, multiply each pixel in the Original Image by the sum of all pixels to the left of it.
- The last pixel in the Integral Image's bottom right corner will equal the total of all the pixels in the Source Image.
- For whatever feature size, the Integral Image requires only four constant value

additions each time (with respect to the 18 additions earlier).

- As the number of adds no longer depends on the number of pixels enclosed, the temporal complexity of each addition rapidly decreases.
- As a result, we employ integral images. The integral image approach simplifies the calculations and saves for any facial detection model, a significant amount of time is required, even if your grid is 1000 x 1000 pixels.

All we must do now is add the purples and subtract the greens from the four corners of our feature.

$$23 = 168 - 114 + 79 - 110.$$

3.1.3 ADABOOST TRAINING

The training phase for the Viola-Jones object detection algorithm use AdaBoost to choose subset of features and create the classifier. A big group of photos is created, the size of which corresponds to the detection window's size. This set must include both positive and negative examples for the chosen filter (e.g., just front view of faces) (non faces). The index of each picture is l , where $l = 1 \dots L$. Each image has a y_l value associated with it. $y_l=1$ for faces and $y_l=0$ for non-faces.

Set the parameters in adaboost-training-eq1 for $y_l=0,1$ and $P-$ and $P+$, respectively, the number of non-faces and faces in the image set.

Run a program for whatever number of rounds you like.

If we run the program for a discretionary number of rounds.

For $i=1$ to l

1. Standardize the loads as follows to make $w_{i,l}$ a likelihood circulation:

$$\frac{w_{i,l}}{\sum_{j=1}^n w_{i,j}} \rightarrow W_{i,l} \dots \dots \dots (1)$$

2. For each element j , train a classifier h_j that can just utilize one element. The classifier's error rate is expressed in terms of $w_{i,l}$:

$$\epsilon_j = \sum_{l=0}^{L-1} w_{i,l} |h_i(x_i) - y_i| \dots (2)$$

3. For each element j , train a classifier h_j that can just utilize one element. The classifier's error rate is expressed in terms of w_i, l :

$$w_{i+1,l} = w_{i,l} \beta_i^{1-e} \dots (3)$$

$$\beta_i = \frac{\epsilon_i}{1-\epsilon_i}$$

4. The final Classifier is:

$$h(x) = \begin{cases} 1, \sum_{i=0}^{L-1} \alpha_i h_i(x_i) \geq \frac{1}{2 \sum_{i=0}^{L-1} \alpha_i} \\ 0 \quad \text{Otherwise} \end{cases}$$

$$\alpha_i = \log \frac{1}{\beta_i} \dots (4)$$

3.1.4 CASCADING CLASSIFIERS

- They introduced the Attentional Cascade technique to make things simpler. The idea is that not every window needs to have all the functionality active. We can conclude that the facial features are absent if a characteristic fails on a certain window. Therefore, we may go on to the following windows, which may include face characteristics.
- The characteristics on the photos are added one at a time. Early phases include simpler features in comparison to later stages, which are more complex, intricate enough to locate the smallest details on the face.
- Remove the window from the remainder of the method if the first step fails to recognise anything on it and move on to the next screen. Because the irrelevant windows will be passed through in most phases, this will save a lot of processing time. Only after the characteristics from the first step are discovered in the image will the second stage processing begin.

3.1.2 FEATURE EXTRACTION

Now that the face has been cropped out of the image, extract certain attributes from it. Look at how to extract these aspects of the face using face embeddings. A neural network receives an image of a person's face as input and provides a vector

that reflects the face's most essential attributes as output. That vector is known as embedding in machine learning, so we'll call it face embedding. Now, how will this be used to recognise different people's looks.

When training the neural network, it learns to output comparable vectors for faces that resemble one other. Consider the following scenario: If I have many photographs of faces from different timelapses, it's evident that some traits will change, but not much.

As a result, the vectors associated with the faces in this problem are comparable, or we could say they are quite close in vector space. Now that we've learned how this network works, let's explore how we can apply it to our own data. We now send all the images in our data to this pre-trained network, which will extract the appropriate embeddings and save them in a file for the next step.

3.2. THINGSPEAK

ThingSpeak is a cloud based IoT analytics tool that allows you to gather, view, and analyse real-time data streams. Collect, Analyze, and Act are the three procedures that it employs. The information we send to ThingSpeak is saved in a channel, which is made up of the following components:

8 data storage fields - These fields can be used to store data from a sensor or an embedded device.

Three location fields can be saved here: latitude, longitude, and elevation.

3.3. WORKING

Hardware components of developed system includes Raspberry-Pi3, Web camera, SD card, Buzzer, MLX90614-IR sensor.

Software requirements included are Python IDE (Any python compiler can be used), Thingspeak.com, NOOBS, VNC viewer.

4. EXPERIMENTAL RESULT

This work captured the user face for nearly 20 samples and stored the data set as user 1 then we trained the 20 faces. After creating the dataset, need to test it by making the system detect any one's face from the dataset. So, created a dataset of entire class students. The system successfully detected our faces and attendance has been marked to our respective roll numbers. Here figure.4 shows the creation of dataset of a student by entering his ID. After this the system will capture 820 images of the respective student and store it under his ID in database as images.

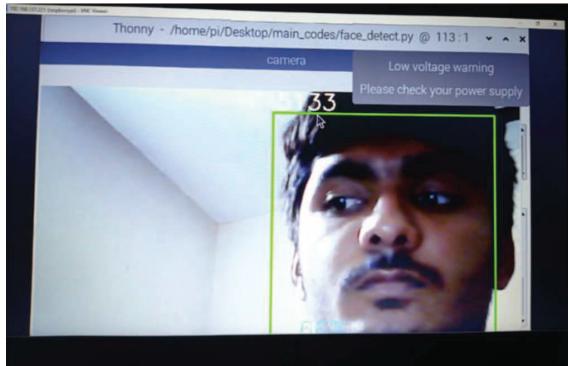


Fig.4. Dataset Creation and Face detection using viola jones algorithm

Figure.4 shows the face detection of the student whose face was detected using viola jones algorithm and is compared to the faces present in the database.

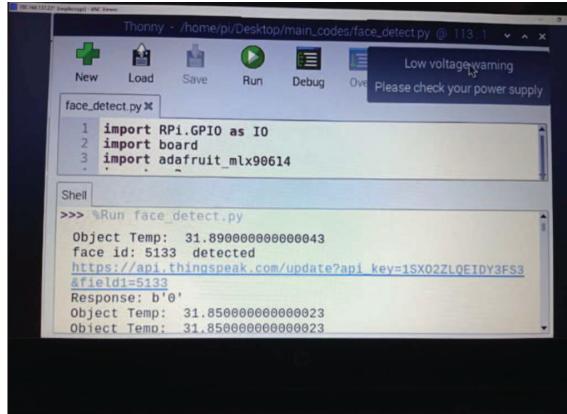


Fig.5. Temperature Check and Face Recognition.

Figure 5 explains about the face recognition rate and correct temperature detection accuracy while the number of faces increases. Images are collected from the online source (YouTube etc).

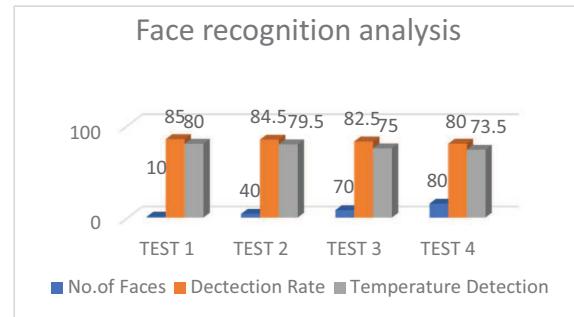


Fig.6.Bar graph for Face recognition analysis.

So, if the number of faces increases the face recognition rate decreases slightly and the temperature detection rate also decreases. Test1, Test2, Test3, Test4 are performed on considering the different number of images the procedures taken for Face recognition and Temperature Detection, by the test results we determine the outcome of the project and accuracy of our product. Thus, our project works efficiently when there are 1-5 maximum faces and temperature detects for the person who is near to the temperature sensor.

Name of the Data sets	Data sets trained (YES/NO)	Face detection	Face Recognition Accuracy
FERET (20 persons)	Yes	88%	80%
Sample faces from online	Yes	78%	73.5%
Real time images from cam	Yes	76%	74%

Table1.Face detection analysis

Initially 10 images taken from FERET dataset of a same person with different angles is used and get the data set trained, when we perform the face detection operation it worked successfully with good efficiency rate. Currently 4GB memory is used. After that same algorithm is implemented on sample images collected from the online sources. Then implemented real time system using raspberry pi and got the performance of 74%.

5. CONCLUSION

We discovered folks employing different ways and procedures for the same task that we were trying to accomplish in our case study. Therefore, we've structured our job to be precise in outcomes and different from others. Our first job was to create a smart attendance tracking system, and we used the above-mentioned components as our technique.

The developed system achieved better results than existing work. For incomplete and dense pictures, the system has currently achieved an accuracy level of up to 80%. It may be developed further to achieve better degrees of accuracy.

In comparison to others, our work receives a speedy response, and the method utilised is extremely suggestible for obtaining reliable results. Furthermore, a low-cost system was constructed; in comparison to similar previous studies, the cost was relatively cheap.

6. FUTURE WORK

Quick reaction time, minimal cost, easy to develop and design, and precise outcomes are some of our project's features. When it comes to developments and future projects, the system takes time to recognise faces when there are numerous persons standing in front of the camera.

Currently, our product is just suitable for temperature checks and attendance. Planning to extend our work on classroom surveillance where number of faces detected can be a student count check in between class. These advancements can make this system better and can be used for massive purposes. Each picture from two or more IP cameras may be handled independently

7. REFERENCES

- [1] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," IEEE Security Privacy, vol. 1, no. 2, pp. 33– 42, Mar./Apr. 2003.
- [2] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," EURASIP J. Adv. Signal Process., vol. 2008, pp. 113–129, Jan. 2008.
- [3] SanjanaPrasad, Mahalakshmi, A. John Clement Sunder R. Swathi" Smart Surveillance Monitoring System Using Raspberry PI and PIR

Sensor" International Journal of Computer Science and Information Technologies (IJCSIT) ISSN 0975-9646 Vol.5 (6), 2014, 7107-7109.

[4]. Sheesha Chintalapati; M. V. Raghunath, "Automated Attendance Management System Based on Face Recognition Algorithms", 2013 IEEE International Conference on Computational Intelligence and Computing Research.

[5]. Phichaya Jaturawat; Manop Phankokkruad, "An Evaluation of Face Recognition Algorithms and Accuracy based on Video in Unconstrained Factors", 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)

Renewable Energy Sources and Electric Vehicle for Optimal Energy Management of Micro Grids with the Aim of CO₂ Emission Reduction

S. Naveen Prakash and N. Kumarappan
Department of Electrical Engineering, Annamalai University, Annamalai Nagar
India
Email: snavenprakash0@gmail.com, kumarappann@gmail.com

Abstract—with increasing issues about global warming and the depletion of fossil fuel sources and there is a need for alternative clean energy, the electrification of vehicles in micro grids has emerged as an undeniable alternative to meet current difficulties. As part of micro grid energy management, the system aims to reduce system purchasing costs while meeting all technological limits and meeting environmental requirements. As a result, this study presents a unique approach for effective charging/discharging schedules of electric vehicles in micro grid in order to reduce network purchasing costs while also lowering Carbon dioxide emission by taking into consideration of different kinds of Generating units (i.e., PV, WT, MT, FC and DG). Under the realistic restrictions of EVs and DERs, the influence of electric vehicle aggregation on operating cost, procuring electricity from main networks and air pollution (CO₂) has been analyzed by using a novel improved whale optimization algorithm. A typical test case is being used to assess the suggested model under two different scenarios. In the assessments, simulation results indicate that the presented approach reduces total cost and carbon dioxide emission significantly. Finally, the obtained results are compared with other optimization techniques.

Keywords—Electric Vehicle, CO₂ Emission, Renewable Energy Sources, Micro grid, Improved Whale Optimization Algorithm

I. INTRODUCTION

The massive rise in energy consumption as a result of increasing and the phenomenal development of industry is one of most serious problems facing the power infrastructure [1]. According to an IEA report, power demand will rise by nearly 50% in the next 25 years, putting a strain on present energy sources [2], [3]. Issues about lowering coal and oil and high air pollution, that is mostly generated by mass transportation and vehicles with internal combustion, have prompted a major growth in the usage of electric mobility and renewable power in microgrids [4]-[6]. Indeed, electric vehicles enable governments to drastically cut carbon emissions in transportation requirements, making them critical to the achievement as part of a worldwide effort to achieve environmental energy commitments and minimize dependency on increasingly limited fossil fuels.

Due to the unpredictability around RES production, a highly desired to increase the saturation of energy storage in order to make such resources dispatchable, the energy grid has become much more complicated as a result of integrating various green sources to each other [7], [8]. Because of the massive carbon dioxide gas emitted by traditional vehicles into the atmosphere, as well as the restriction of expansion in power generation and transmission capacities, electric vehicles have received a lot of attention all around worldwide in these circumstances [9]. The use of electric vehicles instead of gasoline-powered vehicles was proposed

with the goal of increasing the number of electric vehicles on the road to 20 million by 2020 [10]. The battery packs in electric vehicles allows them to keep the electricity they need to travel. These battery packs can be charged through the distribution system. Furthermore, V2G technology allows the distribution system operator to access the stored energy in the battery system [11], [12], so that if the network is unable to meet the demand, they can use the stored energy in the electric vehicle batteries to meet the electricity demand directly. This research proposes a novel cost-emissions-based framework; to solve these issues using improved whale optimization to optimize the operation of EVs when they are coupled with various dispatchable and renewable energy sources, such as fossil fuels and RES. EVs are explored with various travel patterns in order to reduce both procurement costs and air pollution in MGs at the same time.

The remaining of this paper is categorized as follows:

In section 1 describes the introduction of this proposed paper. The problem formulation of the proposed model is presented in section 2. The proposed methodology is discussed in section 3. Results discussion is explained in section 4. Finally, the conclusion is presented in section 5.

II. PROBLEM FORMULATION

The presented multi-objective approach for micro-grid has two objective functions, which are as follows:

$$\text{Multi-objective functions} = \begin{cases} F1 \text{ Cost Minimization} \\ F2 \text{ Emission Minimization} \end{cases}$$

Where, F1 and F2 represent the objective functions presented in details in the following.

$$\text{Cost Minimization (F1)} =$$

$$\text{Min} \left[\text{Cost}^{PV} + \text{Cost}^{WT} + \text{Cost}^{DG} + \text{Cost}^{FC} \right] + \text{Cost}^{Grid} + \text{Cost}^{EV} \quad (1)$$

$$\text{Cost}^{PV} = \sum_{m=1}^{NM} \sum_{t=1}^{NT} [\text{Cost}_{O\&M}^{PV}] = \sum_{m=1}^{NM} \sum_{t=1}^{NT} [K_{O\&M}^{PV} P_{m,t}^{PV}] \quad (2)$$

$$\text{Cost}^{WT} = \sum_{m=1}^{NM} \sum_{t=1}^{NT} [\text{Cost}_{O\&M}^{WT}] = \sum_{m=1}^{NM} \sum_{t=1}^{NT} [K_{O\&M}^{WT} P_{m,t}^{WT}] \quad (3)$$

$$\text{Cost}^{DG} = \sum_{m=1}^{NM} \sum_{t=1}^{NT} [\text{Cost}_{O\&M}^{DG} + \text{Cost}_{fuel}^{DG}] \quad (4)$$

$$\text{Cost}_{fuel}^{DG}(P_{m,t}^{DG}) = k_m I_{m,t} + \Delta T \sum_n^{Nn} \pi_{m,n} P_{m,n,t} \quad (5)$$

$$\text{Cost}_{O\&M}^{DG} = K_{O\&M}^{DG} P_{m,t}^{DG} \quad (6)$$

$$\begin{aligned} \text{Cost}^{FC} &= \sum_{m=1}^{NM} \sum_{t=1}^{NT} [\text{Cost}_{fuel}^{FC} + \text{Cost}_{O\&M}^{FC}] \\ &= \sum_{m=1}^{NM} \sum_{t=1}^{NT} \left[\left(\frac{C_{ng}}{L_{ng}} \frac{P_{m,t}^{FC}}{\eta_{m,t}^{FC}} \right) K_{O\&M}^{FC} + P_{m,t}^{FC} \right] \quad (7) \end{aligned}$$

$$\begin{aligned} Cost^{MT} &= \sum_{m=1}^{NM} \sum_{t=1}^{NT} [Cost_{fuel}^{MT} + Cost_{O\&M}^{MT}] \\ &= \sum_{m=1}^{NM} \sum_{t=1}^{NT} \left[\left(\frac{C_{ng}}{L_{ng}} \frac{P_{m,t}^{MT}}{\eta_{m,t}^{MT}} \right) K_{O\&M}^{MT} + P_{m,t}^{MT} \right] \end{aligned} \quad (8)$$

$$Cost^{Grid} = \sum_{t=1}^{NT} C_t^{Grid} P_t^{Grid} \quad (9)$$

$$\begin{aligned} Cost^{EV} &= \left| \sum_{t=1}^T (EV_{load}(t) + P_{EV}(t)) \right| \\ (s.t. P_{EV}(t) \leq 0, EV_{load}(t) \leq 0) \quad (10) \end{aligned}$$

$$Emission Minimization (F2) = \sum_{i=1}^N (\rho_i \times P_{m,t}) \quad (11)$$

$$Emission Cost (C_{EC}) = \sum_{j=1}^3 \gamma_j \times (\sum_{i=1}^N \rho_i \times P_{m,t}) \quad (11.a)$$

A. EV constraints

The below mentioned constraints are imposed on the operation limit of electric vehicle.

$$EV_m^{Min} \leq EV_{m,t} \leq EV_m^{Max} \quad (12)$$

Where, the maximum and minimum power levels of electric vehicles are mentioned in (12).

B. Power balance constraints

The total real power generated by generation units plus power purchased from utility grid must balance the predicted power demand of the system plus the sold power to utility grid, at time t .

$$\begin{aligned} P_{load,t} + \sum_t P_{sell,t} = \\ P_t^{Grid} + \sum_{m=1}^{NM} P_{m,t}^{PV} + P_{m,t}^{WT} + P_{m,t}^{FC} + P_{m,t}^{DG} + P_{m,t}^{Disch} \end{aligned} \quad (13)$$

$$-P_{Grid,max} \leq P_t^{Grid} \leq P_{Grid,max} \quad (14)$$

III. PROPOSED METHODOLOGY

This proposed method, improved whale optimization algorithm [13], is hybridized the whale optimization algorithms operators with differential evolutions mutation operator to improve the capability of exploration and exploitation of the proposed method. The major operators are hybrid operators, which integrate DEs modification with whale optimization algorithm features such as surrounding prey, searching for prey, and spiral updating location. IWOA is divided into two sections: exploration and exploitation. The exploring part modifies the individuals when rand is less than lambda. It is used to alter it by Equation (15).

$$\lambda = 1 - t/t_{max} \quad (15)$$

Where t denotes the current generation and t_{max} denotes the total number of generations. Lambda is utilized to control IWOA's exploration and exploitation abilities. Between one and zero, it can be gradually reduced. As a result, individuals are free to analyze the early generations. As time passes, the exploitation continues. DE's evolution and the seeking for WOA sustenance are found in the exploration section. However, the presented technique's exploitation portion is comparable to WOA. Note that boundary constraints must be considered when problem solutions are presented. If any of these limitations are exceeded, the restoring procedure is continues to follow Equation (16).

$$X_i(j) = \begin{cases} \delta_j + rndreal(0,1) \times (\mu_j - \delta_j) & \text{if } X_i(j) < \delta_j \\ \mu_j + rndreal(0,1) \times (\mu_j - \delta_j) & \text{if } X_i(j) < \mu_j \end{cases} \quad (16)$$

IV. RESULTS AND DISCUSSIONS

A. Mode 1

In this mode has operated and satisfied the electricity demand without presents of renewables and electric vehicle. It has been operated in connected mode. The objective is to find the compromised optimal scheduling of the proposed units with lowest cost and reduced carbon dioxide emission. The scheduling of this test system has considered into 24-h time horizon. It is also noted that using other sources, such as diesel generators leads to increasing in pollution, due to their highest emission co-efficient; while it causes high operation cost because the bids of units are high, i.e. in the economic point of view, employing these kinds of sources must be limited according to its economic considerations. Because the availability of renewable energy sources is reducing, the amount of energy consumed from the main grid is increasing. Furthermore, deficiencies of renewable energy sources are replaced through other distributed energy units and the main grid resulting in higher DG unit generating prices as well as higher system Carbon dioxide emission in environment. Table.1 shows the numerical results obtained in mode 1 operation.

B. Mode 2

The effective energy planning of micro grid is addressed in this research to reduce both cost and Carbon dioxide emissions through effective scheduling and inclusion of electric vehicles and renewables respectively. Furthermore, by shifting power use away from maximum demand hours to more desired hours when electricity need has stabilized the electric vehicles can lower the system operation cost. In contrast to this change, research has shown that electric vehicles can significantly reduce Carbon dioxide emissions. The numerical results obtained in this suggested model are provided and discussed in details. Table.2 and Table.3 illustrate the amount of active power produced by distributed generation units and electric power purchased from the utility grid as well as the amount of generation with electric vehicle and renewables respectively. The presence of electric vehicle on the network was analyzed, as shown in Table.3. The results show that in the presence of electric vehicle, the power purchased from the utility system is not only lowered, but also reduced with the use of vehicle to grid technology and the purchase of energy saved in electric vehicle.

Furthermore, according to Table.3 Carbon dioxide emissions are reduced by minimizing energy consumed as from utility system and traditional DG output. The diminution of atmospheric carbon dioxide in the availability of electric vehicle is depicted in Figure.1. Due to higher power consumption during the day's peak hours (hours 1-6 and 19-24), EV has been released during such hours to ensure reliability and stability and contribute to the network's long-term sustainability and reliability. These vehicles, on the other hand, were capable of charging their cells during off-peak hours (7-18) using EVCC programme. In terms of the cost-minimizing goal, however, electric

vehicle discharging takes place during peak times when energy prices are high.

IWOA was utilized to optimize micro grid management with the aim of decreasing overall DG cost and reducing Carbon dioxide emission. In Table.4, the results obtained from this proposed optimization techniques are provided in comparison with other optimization techniques such as whale optimization algorithm (WOA) and genetic algorithm (GA). As per the results, it is concluded both the total cost of micro grid and Carbon dioxide emission have decreased significantly. Figure.2 and Figure.3 shows the convergence properties of mode 1and mode 2 of operations respectively.

TABLE I. OPTIMAL SCHEDULING OF DG UNITS (MODE 1)

Time (h)	OUTPUT POWER (kW)				
	FC	DG1	DG2	MT	Utility
1	152.13	22.62	15.68	105.57	-15.00
2	69.05	39.87	63.17	122.87	-29.96
3	127.02	31.09	15.33	157.06	-30.50
4	135.86	35.69	22.92	108.64	-28.12
5	132.69	23.22	14.54	53.43	-25.88
6	96.40	57.45	14.26	71.89	-30.00
7	115.93	16.67	10.70	73.03	11.67
8	37.32	23.67	24.42	85.22	38.37
9	121.51	24.90	36.80	66.77	24.01
10	69.30	27.42	38.75	112.64	46.89
11	82.10	10.02	27.51	58.35	14.02
12	57.35	14.93	10.25	69.19	13.28
13	60.91	26.82	17.72	65.55	22.00
14	36.74	21.07	10.59	70.39	31.21
15	40.21	11.44	19.81	55.17	38.37
16	63.86	10.31	13.75	70.62	41.46
17	57.74	20.15	36.18	91.15	15.78
18	91.75	37.35	28.10	102.36	10.44
19	88.28	27.58	21.72	93.46	-21.04
20	95.95	17.07	26.38	52.56	-11.96
21	68.28	12.20	14.95	91.19	-21.62
22	88.19	32.35	36.20	67.30	-21.05
23	50.90	24.66	28.02	105.71	-19.28
24	81.40	10.18	29.33	102.69	-28.60

TABLE II. OPTIMAL SCHEDULING OF DG UNITS (MODE 2)

Time (h)	OUTPUT POWER (kW)			
	FC	DG1	DG2	MT
1	226.27	38.52	10.18	107.32
2	168.34	28.35	38.29	105.75
3	198.18	53.35	32.57	97.08
4	218.97	18.40	21.50	104.01
5	163.69	20.15	16.89	80.40
6	158.65	13.95	11.64	125.24
7	33.87	10.60	13.25	52.64
8	43.39	11.38	11.40	52.59
9	123.50	10.09	10.35	56.42
10	74.58	29.20	33.11	74.34
11	71.07	11.56	11.55	56.84
12	40.09	10.19	10.24	50.06
13	36.37	10.40	10.05	51.49
14	30.74	11.44	10.05	51.13
15	35.88	10.48	11.27	50.56
16	52.01	13.09	12.83	76.22
17	94.58	11.86	10.31	53.43
18	53.93	33.18	24.12	56.53
19	151.74	28.44	26.47	104.47
20	110.03	12.46	13.83	75.67
21	99.55	13.06	12.23	72.64
22	132.33	12.58	11.95	73.09
23	159.44	32.80	38.51	70.63
24	193.94	18.57	13.58	68.33

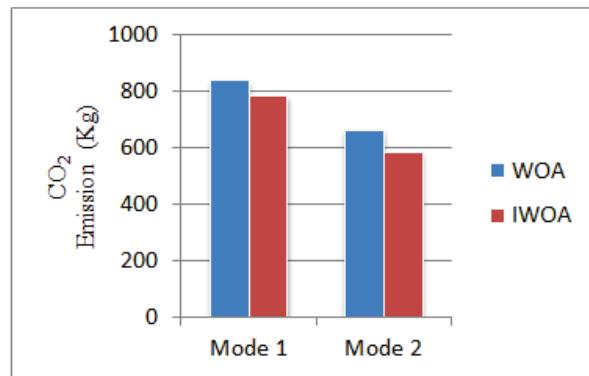


Fig. 1. Comparison of CO₂ emission.

TABLE III. OPTIMAL SCHEDULING OF RENEWABLES, EV AND UTILITY (MODE 2)

Time (h)	OUTPUT POWER (kW)			
	PV	WT	Grid	EV
1	0	1.7	-22	-81
2	0	8.5	-4.22	-80
3	0	9.27	-10.45	-80
4	0	16.66	-23.55	-81
5	0	7.22	-4.36	-86
6	0.03	4.91	-11.43	-93
7	6.27	14.66	20.71	96
8	16.18	25.56	13.5	35
9	24.05	20.58	14.21	14.79
10	39.37	17.85	13.52	13.03
11	7.41	12.8	0	20.77
12	3.65	18.65	5.7	26.41
13	31.94	14.35	16.48	21.92
14	26.81	10.35	33.1	20.38
15	10.08	8.26	22.67	15.79
16	5.3	13.71	16.83	10
17	9.57	3.44	17.8	20
18	2.31	1.87	18.06	80
19	0	0.75	-11.87	-90
20	0	0.17	-11.16	-21
21	0	0.15	-21.2	-11.43
22	0	0.31	-17.26	-10
23	0	1.07	-19.45	-93
24	0	0.58	-15	-85

TABLE IV. COMPARISON OF NUMERICAL RESULTS

Optimization Techniques	Modes of Operation			
	Mode 1		Mode 2	
	Cost (\$)	Emission (Kg)	Cost (\$)	Emission (Kg)
IWOA	3193	787.04	2952.73	586.79
WOA	3234	841.92	3122.62	667.88
GA	3261	857.86	3336.19	722.67

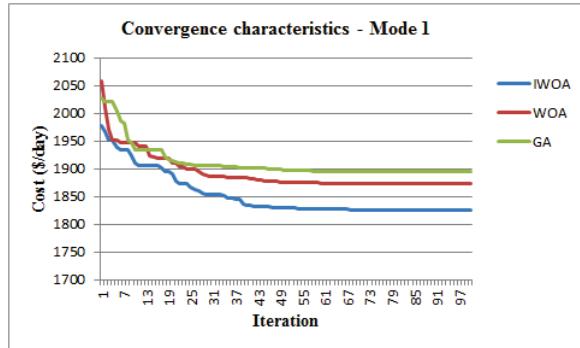


Fig. 2. Convergence characteristics – Mode 1

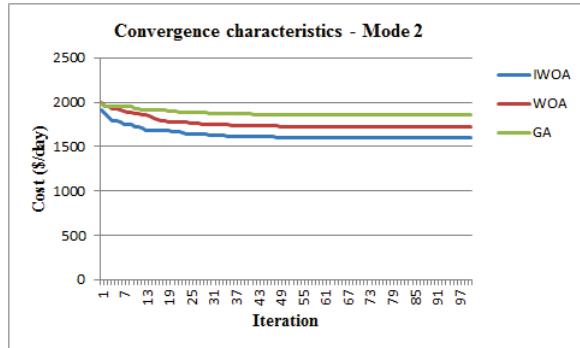


Fig. 3. Convergence characteristics – Mode 2

V. CONCLUSION

In this study, an optimal power scheduling framework for DG units and EVs was extended based on cost-emission models. In the proposed model, several DG units are considered, such as PV, FC, MT, WT, and DG. In this optimization problem, the main objectives are to minimize the total MG cost, to reduce carbon dioxide emissions and to consider the EV and DER units' technical constraints. Considering that the problem is formulated as an optimization framework with time constraints, IWOA optimization is applied to solve it. From the results it can be concluded, there was a significant reduction of total cost of operation for DG resources. The presented approach was tested using MG test system. The overall cost and DG operational costs were reduced significantly in the study. According to the numerical results, using electric vehicle and renewables can result in significant CO₂ emission reductions at the MG. Including the emission objective function in simulations resulted in modifications to the charging/discharging schedule to reduce the total air pollution emissions. In addition, the comparisons of convergence of the proposed improved WOA algorithm show its greater effectiveness when compared to other optimization techniques.

Nomenclature

Abbreviation

DG Diesel generation units

FC Fuel cell

MG Micro grid

MT	Micro-turbine	$P_{load,t}$	The total load of Micro grid at time t (kW)
PV	Photovoltaic	$P_{m,t}^{PV} P_{m,t}^{WT}$	The Generated power of Photovoltaic and Wind Turbine at time t (kW)
RES	Renewable energy sources	$P_{m,t}^{FC} P_{m,t}^{MT}$	The Generated output power of Fuel Cell and Micro Turbine at time t (kW)
WT	Wind turbine	$P_{m,t}^{DG}$	The Generated output power of Distributed Generation units (kW)
<i>Indices</i>		$\eta_{m,t}^{FC} \eta_{m,t}^{MT}$	The Electrical efficiency of Fuel Cell & Micro Turbine (%)
m	Index of MGs		
t	Index of time		
$Cost_t^{Grid}$	Electricity price of grid at time t (\$/kWh)		
$K_{O\&M}^{PV} K_{O\&M}^{WT}$	Constant coefficient for O&M cost Photovoltaic and Wind Turbines (\$/kW)		
$K_{O\&M}^{DG}$	Constant coefficient for O&M cost diesel generation units (\$/kW)		
$K_{O\&M}^{FC} K_{O\&M}^{MT}$	Constant coefficient for O&M cost Fuel Cell and Micro-Turbine (\$/kW)		
$P_{m,t}^{FC} P_{m,t}^{MT}$	Installed capacity of Fuel Cell and Micro-Turbine (kW)		
$P_{grid,max}$	Maximum power purchased from utility grid (kW)		
$P_{m,t}^{WT}$	Installed capacity of WTs (kW)		
$Cost^{PV}$ $Cost^{WT}$	Overall cost of PVs and WTs (\$)		
$Cost_{O\&M}^{PV}$ $Cost_{O\&M}^{WT}$	O&M cost of Photovoltaic and Wind Turbine (\$)		
$Cost^{FC}$ $Cost^{MT}$	Overall cost of Fuel Cell and Micro-Turbine (\$)		
$Cost_{fuel}^{FC}$ $Cost_{O\&M}^{FC}$	Fuel cost and O&M cost of Fuel Cell (\$)		
$Cost_{fuel}^{MT}$ $Cost_{O\&M}^{MT}$	Fuel cost and O&M cost of Micro-Turbine (\$)		
$Cost^{DG}$	Overall cost of Distributed Generation units (\$)		
$Cost_{fuel}^{DG}$ $Cost_{O\&M}^{DG}$	Fuel cost and O&M cost of Distributed Generation units (\$)		
$Cost^{Grid}$	Purchased cost of power from utility grid (\$)		
$F1$	Objective function- Cost		
$F2$	Objective function- Emission		
P_t^{Grid}	Electrical Power purchased from utility grid at time t (kW)		

REFERENCES

- [1] S. M Dawoud, X Lin, M.I Okbam, "Optimal placement of different types of RDGs based on maximization of microgrid loadability," Journal of Cleaner Production, vol. 168, pp. 63-73, 2017.
- [2] OECD/IEA, OECD Green Growth Studies Energy. [Online]. Available:<http://www.oecd.org/greengrowth/greeningenergy/49157219.pdf>, Accessed May 2011.
- [3] M.Z. Jacobson, M.A Delucchi, "Providing all global energy with wind, water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials," Energy policy, vol. 39, pp. 1154-1169.
- [4] K. Laurischkat, D. Jandt, "Techno-economic analysis of sustainable mobility and energy solutions consisting of electric vehicles, photovoltaic systems and battery storages," Journal of Cleaner Production, vol. 179, pp. 642-661, 2017.
- [5] M. Montazeri, K. Mahmoodi, "Optimized predictive energy management of plug-in hybrid electric vehicle based on traffic condition," Journal of cleaner production, vol. 139, pp. 935-948, 2016.
- [6] R.Á. Fernández, "A more realistic approach to electric vehicle contribution to greenhouse gas emissions in the city," Journal of Cleaner Production, vol. 172, pp. 949-959, 2018.
- [7] F.S. Gajizahani, J. Salehi, "Stochastic multi-objective framework for optimal dynamic planning of interconnected microgrids," IET Renewable Power Generation, vol. 11(14), pp. 1749-1759, 2017.
- [8] F.S. Gajizahani, J. Salehi, "Robust design of microgrids with reconfigurable topology under severe uncertainty," IEEE Transactions on Sustainable Energy, vol. 9(2), pp. 559-569, 2017.
- [9] P. Aliasghari, B. Mohammadi-Ivatloo, "Optimal scheduling of plug-in electric vehicles and renewable micro-grid in energy and reserve markets considering demand response program," Journal of Cleaner Production, vol. 186, pp. 293-303, 2018.
- [10] M. Ghofrani, A. Arabali , M. Etezadi-Amoli, "Smart scheduling and cost-benefit analysis of gridenabled electric vehicles for wind power integration," IEEE Transactions on Smart Grid, vol. 5(5), pp. 2306-2313, 2014.
- [11] H. Xing, M. Fu, Z. Lin, Y. Mou, "Decentralized optimal scheduling for charging and discharging of plug-In electric vehicles in smart grids," IEEE Transactions on Power Systems, vol. 31(5), pp. 4118-4127, 2016.
- [12] J.R. Pillai, B. Bak-Jensen, "Integration of vehicle-to-grid in the western Danish power system," IEEE Transactions on Sustainable Energy, vol. 2(1), pp.12-19, 2011.
- [13] Samaneh Yazdani and Seyed Mostafa Bozorgi, "IWOA: An improved whale optimization algorithm for optimization problems," Journal of Computational Design and Engineering, vol. (6), pp. 243-59, 2019.

Author Index

- Allirani, S. 168
Akash Ram R. K. 329
Akhilesh Tyagi 1
Akila, V. 91
Akshya. J. 196
Akshay, S. 52
Ampapurapu, Mounika Durga 85
Anandan, P 111
Anis, Md Tarique 191
Archana, R. 23
Arumugam, Sajeev Ram 158

Babu, N. 117
Balakrishna, R. 158
Baranidharan, B 101
Bharath, Banovath Shiva Prasad 206
Bharathiraja, N. 247
Bhargavi, Setlem 46
Boovaraghavan, Aravinda 123
Bushra, S Nikkath 185, 357

Chakrapani, Praveena 335, 341
Chang, Chih-Yung 13, 18
Chansekar, S. 259
Charaan, S. 276
Chelliah, Balika J 225
Chitradevi, D. 335, 341
Christy Jackson, J. 123

Dande, Bhargavi 13, 18
Daniel Einarson 7
Dawit Mengistu 7
Deepa, K. 46
Deepthi, V. Hima 382
Deepika Shetty M. A. 52
Devi, E. Anna 158
Dhasny Lydia, M. 299
Dinesh, P. 259
Domakon, Srikanth 263
Doss, Rajesh 35
Durga Kameswari, S. 241

Elavarasan, P. 276
Ezhilmathi, S. 310

Fan, Cheng-De 13, 18
Fatima, N. Sabiyath 364
Felix, A 64

Gadupudi, Prudhvi 196
Ganesh Kumar, C. 329
Gayathri, R 287
Gayathri, S. 351
Godbin, A Beena 185
Gunasekar, T 64
Gurijala Sreedhar 212

Hari Prasath, B. 259
Hema, A. 107
Hemalatha, K. 140
Hashwanth, S. 123

Indiramma, M. 253

Jagannaveen, V. 241
Jamal, K. 263
Janet, B. 79
Jayashree, K. 370
Jyothirmai, G. 75
Jyothsna, Uppuluri Baby 46

Kamaleshwar, T 305
Kannan, S. Gokul 91
Karthik, S. 64
Karthick, T. 128
Karuppasamy, Sankar Ganesh 158
Kasthuri, P. 276
Kavitha, A. 91
Kavitha, M 111, 281, 287
Kavuri, Karunakar 281
Kirthiga, M. 163
Kollu, Praveen Kumar 85
Kona, Swathi 152
Konathala, Rakesh 152
Koniki, Rachana 85
Krishna, K. Venkata 75

- Krishnakishore, K. 241
 Krishnamoorthy, R. 247
 Kujani, T. 91
 Kumar, K. Hemanth 75
 Kumar, K. Kiran 107
 Kumar, Rayudu Bharath 235
 Kumar, S. Vinod 158
 Kumaran, Kaushik 101
 Kumarappan, N. 388
 Logesh, S. 259
 Madhumithaa, B 168
 Madhuri, J. 253
 Maheshwari, K. Uma 357
 Maheysh, V. 123
 Manikandan, A 259
 Manjula, P. 58
 Manjula, S 111
 Manoj, R. 219
 Manoj, S 97
 Maria, Azees 152
 Mavuru, Hemanth Kumar 152
 Ming-Yang, Su 13
 Mohammad, Ashraf 41
 Morthala, Sai Venkata Krishna Reddy 152
 Muralikrishna, Ch 235
 Nandakumar, B. 219
 Nalinkumar, S. 276
 Naren G. O. 123
 Narennathan. S. K. 270
 Naveenkuamr, M. 219
 Nayagam, R. Deiva 247
 Neppalli, Reshma 107
 Omana, J. 247
 Padmakar, Ganta Sandeep V 235
 Paladugula, Karthik 196
 Pawar, Sheetal Umakant 173
 Pedapalli, Saam Prasanth Dheeraj 41
 Pinninti, Pawan Kumar 152
 Pooja, K. 317
 Prabavathi, R. 225
 Prakash, M. 299
 Prakash, P. 276
 Prakash, S. Naveen 388
 Prakash, V. S. 247
 Prashant 323
 Prathibha, Soma 163
 Praveenkumar, S. 128
 Pravin Kumar, P. 163
 Premesar, Panditi 382
 Priya, E. 351
 Priya, Sbaghavathi 58
 Priyadarsini, P. L. K. 196
 Rahul, R. 219
 Rajalakshmi, D 305
 Rajalakshmi, N R 292
 Rajenn, Babu 370
 Rajeshram, V. 158
 Rajkamal, M. 305
 Ramadoss, R. 69
 Ramanjaneyulu, K. 75
 Ramesh, N. V. K 107
 Ramprabhakar, J. 46
 Ramya, S. 146
 Rao, Keshav 101
 Raveen, Muddineni 41
 Ravikumar, S. 376
 Reddy, B. Naresh Kumar 107
 Reddy, Gummala Sainath 382
 Revathi, M 185
 Rozen Berg, D. 179
 Rp, Jugi Rithanya 317
 Rubini, B 97
 Sakthivel, R. 317
 Sakthivel, S 329
 Sailaja, V. 46
 Samiuddin, Syed 263
 Samraj, M. Jerome 347
 Sanjay Kannan, M. 179
 Sankar, Adhith 101
 Saranya, S. S. 364
 Saravanan, K 292
 Sarwar, Md. 323
 Satapathy, Er Jeetamitra 235
 Sathishkumar, R. 168
 Sathish, Suriyakrishnan 123
 Selvaraj, A. 64
 Selvakumara Samy, S. 310
 Shabika Fathima, M. 163
 Shalin, N. 347
 Shankar, A. 270
 Shanmuga Priya, R. 179
 Sharma, Veena 191
 Shivani, Pathi 263
 Shobana, G. 357
 Shyamkumar, M. 163
 Sibi, S Anslam 185
 Siddiqui, Anwar Shhzad 323

- Sidarth Sai, B. 329
Singh, Sunita 235
Sivarajanji, P. 376
Snehith, K. 75
Sornalakshmi, K. 347
Souri, Gorla Sujan 382
Sree Dhviya, M. 168
Sree, Manne Navya 200
Sree, Mupparaju Kavya 41
Srinidhi, S. 163
Srinivasan, Makesh 123
Subramaniam, Nalini 357
Suganthy, M. 111
Suguna, N 235
Sumanth, Pabba 263
Sumathy, M 305
Supriya, Ch. Gowthami 107
Suresh, Tamilarasi 117
Sureshkumar, T. N. 270

Thareja, Hitesh 229
Tharoor, Vishnusai Viswajith 134

Tharunraj, M. 179
Thiagarajan, R. 247
Thiyagu, M. 270

Uganya 305
Uma, M. 146

Valluru, Sudarshan K. 229
Vanam, Satyanarayana 200, 206, 212
Varghese Mathew Vaidyan 1
Venkataraman, Revathi 347
Vetriselvi, V 140
Vijay, K. 370, 376
Vijayakumar, R. 370, 376
Vijayaraj, N 305
Vimala, C. 69
Viveka, M. 347

Yashwanth Krishnan, B. 163
Yaswanth Pavan Kumar, T. 41



Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology

Contact Us

No.42, Avadi-Vel Tech Road,
Vel Nagar, Avadi, Chennai – 600 062,
Tamil Nadu, India.
Phone : 044-4306 6864

Join our Community

 hodcse@veltech.edu.in

 www.veltech.edu.in

 <https://www.facebook.com/veltechinstitute/>

 https://twitter.com/VT_Univ

 <https://www.linkedin.com/in/vel-tech-164258a8/>

 https://www.instagram.com/veltech_institute/